DOCUMENT RESUME

ED 080 574 TM 003 090

AUTHOR	Haak, Ruth Adlof; And Others
TITLE	Student Evaluation of Teacher Instrument, II. Manual.
INSTITUTION	Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Research and Development Centers Branch.
BUREAU NO	BR-5-0249
PUB DATE	72
CONTRACT	OEC-6-10-108
NOTE	113p.
AVAILABLE FROM	The Dissemination Department, The Research and Development Center for Teacher Education, The University of Texas, Austin, Texas 78712
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Computer Programs; *Elementary Grades; Factor Analysis; *Manuals; Scoring; *Student Opinion; *Teacher Evaluation; *Teacher Rating; Test Construction; Testing; Test Reliability

ABSTRACT

The Student Evaluation of Teacher II (SET II), the new instrument presented in this manual, is designed for use in securing student evaluations of teachers below the fourth grade or, with disadvantaged students, through the sixth grade. The manual describes the history, nature, and difficulties of the student evaluation of teachers movement and provides all necessary instructions for processing SET II data in any school setting, including those without access to a computer and/or personnel with professional expertise in statistics. The manual consists of the following parts: (1) the rating of teachers by students below high school level--reasonableness, methodology, and results; (2) development of the SET II; (3) the SET II--reliability study, factor analyses, categorical structure (3 factors--stimulating, interactive style, unreasonable negativity, and fosterance of self-esteem), and final form of the instrument; and (4) using the SET II--for feedback and research. The SET 2 Computer Program Manual is attached. (Author/KM)

FILMED FROM BEST AVAILABLE COPY

MANUAL

STUDENT EVALUATION OF TEACHER INSTRUMENT II

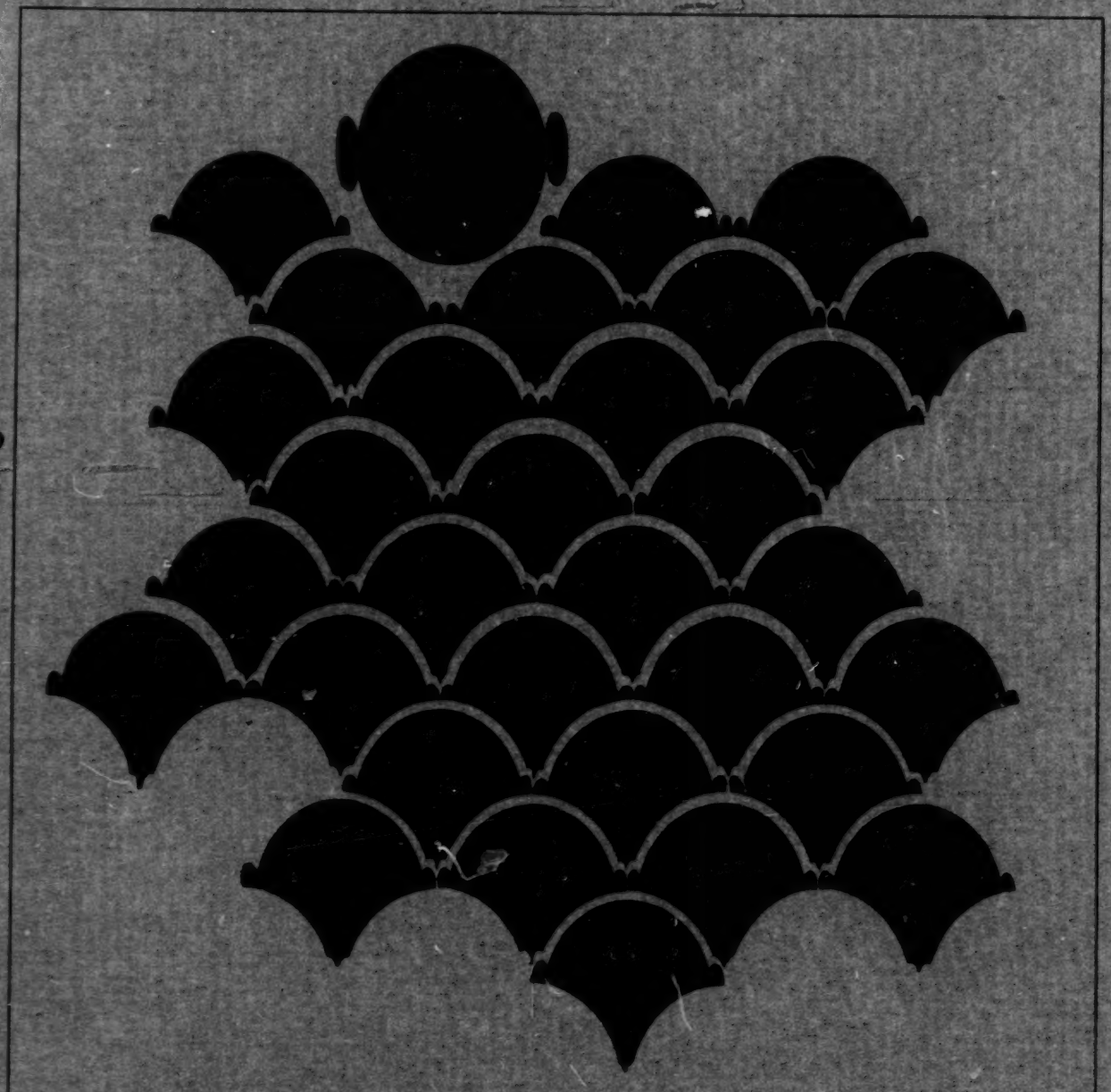
U.S. DEPARTMENT OF HEALTH EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

Ruth Haak
Douglas Kleiber
Robert Peck

whole book

ED 080574

TM 003 090



UTR&D The Research and Development Center for Teacher Education • The University of Texas at Austin • Austin, Texas

ED 080574

STUDENT EVALUATION OF TEACHER INSTRUMENT, II.[1]

MANUAL

Ruth Adlof Haak, Ph.D.

Douglas A. Kleiber, Ph.D.

Robert F. Peck, Ph.D.

The Research and Development Center[2]
for Teacher Education

The University of Texas at Austin
Austin, Texas

1972

Footnotes to Title Page

[1]The instrument described in this manual was developed by the Research and Development Center for Teacher Education under USOE Grant No. OE-6-10-108.

Many persons made important contributions to this work. The need for the instrument was seen by Dr. Robert Peck. Nancy Hinojosa carefully supervised the manuscript through all its various stages. Dr. Don Veldman offered helpful technical advice at many points. Elaine Morris and Kay Carlyle worked far beyond any reasonable level of duty in collecting the data. Louise Jones carefully processed the data for computation. Also, many teachers and students cheerfully cooperated in this project. The cooperation of the administration of the Austin Independent School District was also most gratifying. To all these people and to many unnamed individuals who made contributions to the study the authors wish to express their sincere thanks.

[2]Requests for reprints should be addressed to The Dissemination Department, The Research and Development Center for Teacher Education, The University of Texas, Austin, Texas 78712.

TABLE OF CONTENTS

		Page
I.	Introduction	1
	Special Note	2
II.	The Rating of Teachers By Students Below High School Level: Reasonableness, Methodology, and Results	4
	Review of the Literature	4
	Reasonableness of Doing this Research	7
	Can Young Students Discriminate?	7
	Are Ratings of Young Students Reliable?	11
	Are Ratings of Young Students Valid?	12
	Methodology	13
	Results	18
	Basic Dimensions of Teacher Behavior	18
	A Major Teacher Attitude Gestalt Which Appears to Relate to Student-Identified "Warmth and Friendliness"	27
III.	Development of the Student Evaluation of Teacher (SET II)	30
IV.	The Student Evaluation of Teacher II (SET II)	35
	Reliability Study	36
	Factor Analyses	46
	The Categorical Structure of the SET II	49
	Factor I: Stimulating, Interactive Style	58
	Factor II: Unreasonable Negativity	60
	Factor III: Fosterance of Self-Esteem	62
	Final Form of the Experimental SET II	63
V.	Using the SET II	67
	Using the SET II for Feedback Purposes	68
	Use of the SET II in Research	75
Appendix A.	Individual Item Cards	77
Appendix B.	Sorting Envelope	78
Appendix C.	Available Factor Analyses	79

TABLE OF CONTENTS (cont.)

		Page
Appendix D.	The Factor Structure of the Set II	80
Appendix E.	Teacher Tally Sheet	81
Appendix F.	Teacher Tally Sheet Example	83
Appendix G.	Teacher Profile	85
Appendix H.	Teacher Profile Example	86
Appendix I.	SET II Pupil's Scoring Record	87
Appendix J.	SET II Pupil's Scoring Record Example	89
Appendix K.	Computing Norms	91
Appendix L.	One-Page Form to Use with Grades 4-6	96
Appendix M.	A Comparison of Six Teachers' Profiles	97
References		98

LIST OF TABLES

Table		Page
1	Methods of Eliciting Student Ratings of Teachers	15
2	Comparison of Factors Identified	24
3	Percentage of Agreement Pre to Post: Total Sample at Each Grade Level	37
4	Children in Kindergarten, Percentage of Agreement Pre to Post	39
5	Children in Grades 1-3, Percentage of Agreement Pre to Post	40
6	Children in Grade 4-6, Percentage of Agreement Pre to Post	41
7	Children in Kindergarten, Ability Level Within Race, Percentage of Agreement Pre to Post	42
8	Children in Grades 1-3, Ability Level Within Race, Percentage of Agreement Pre to Post	44
9	Factor Analysis of SET II, Grades 1-3	48
10	Primary (Highest) Loadings, SET II Items---Three Factor Structure, Grades 1-3	50
11	Factor Analysis of SET II, Grades 1-6	52
12	Primary (Highest) Loadings of SET II Items--Four Factor Structure, Grades 1-6	53
13	Comparison of Primary Factor Structures, Derived from Teacher Mean Scores	56

LIST OF TABLES (cont.)

Table		Page
14	Means and Standard Deviations: Pooled Student Responses All Items, Grades 1-3 and Grades 4-6	64
15	Standardized Means: The Teacher Item Means All 36 Teachers, All 22 Items	65
16	Standardized Scales, 20 Item SET II, 36 Teachers	69

CHAPTER I

INTRODUCTION

Asking students to evaluate their teachers' effectiveness is both an old idea and a logical one. Systematic evaluation of teachers by their students is rather widespread at the upper elementary and secondary levels in school. But such evaluations have not been easy to obtain at the lower levels of schooling, i.e., the primary grades 1-3.

The new instrument presented in this manual is designed for use in securing student evaluations of teachers below the fourth grade level. This instrument is also thought appropriate for use with disadvantaged populations through sixth grade. The instrument is named the Student Evaluation of Teacher II (SET II) after its predecessor, the SET I, which is used with older children, grades 5 and above (Veldman, 1970). Both instruments were developed at The Research and Development Center for Teacher Education; University of Texas; Austin, Texas, 78712.

This manual is larger than the typical instrument manual because it represents a certain stance in educational measurements, namely, the stance that any educator who wishes to measure another person for whatever reason should be vitally

concerned about the nature of the measurement that he is hoping to take. Psychological measurement has not reached the state of perfection that the thermometer represents in medicine. Most psychological "measures" are quite imperfect, even when they are a distinct improvement over nothing at all. Therefore a person who wishes simply to measure another person without bothering himself about the nature of the characteristic he is hoping to "measure" is not being realistic about psychological measurement. He will undoubtedly be deluded by his results, and he may inadvertently damage his cause, and other persons, more than he helps either.

This manual attempts to "let the reader in" on the history, nature, and difficulties of the student evaluation of teachers movement. Persons who are interested in the pursuit of student evaluation of teachers are invited to enter into the total complexity of this area. Only with the cooperation of many interested, concerned persons communicating with each other can viable methodologies be developed for understanding what it is that students can tell us about teaching and its effects upon them.

Special Note

This manual is complete with the necessary instructions for processing SET II data in any school setting, including those

which do not have access to a computer and/or lack personnel with professional expertise in statistics. Where these limiting conditions do not prevail, SET II results are, of course, to be computed with standard technological processes.

CHAPTER II

THE RATING OF TEACHERS BY STUDENTS BELOW HIGH SCHOOL LEVEL: REASONABLENESS, METHODOLOGY, AND RESULTS

Review of the Literature

The great length of time that student rating of teachers had been followed as a practice in this country is surprising. In 1896, Kratz obtained the responses of 2,411 elementary and secondary school children to a set of structured questions about teaching. He found, not too unlike 1970, that children liked teachers who were encouraging, patient, polite, neat, and pleasant (Kratz, 1896). The area of student rating of teachers subsequently became so active that the 1931 Review of Educational Research devoted a whole chapter to "Teacher Rating," a chapter followed by 83 references.

Today, however, the question of "What makes a good teacher?" still lingers on. Partly because of the cited age and activity of this line of research, its continuance into the present is not easy to understand. For example, Weintraub points out that research in this area is remarkable for its lack of sophistication, but the sheer redundancy of the findings leave the answer hardly in doubt. Over and over the results of

this line of inquiry indicate that students "want to be understood, that they want a sympathetic teacher, that they want one with a sense of humor, and that they want a teacher who can teach ". (Weintraub, 1967, p. 443).

There are several reasons for the fact that research into students' perceptions of teachers still continues. Weintraub himself points out the most obvious one: "...the statements are not so simple as they may seem at first glance. An operational definition of an understanding teacher might be rather complex. Indeed, the concepts of understanding and sympathy are probably interpreted differently by different individuals (p. 443)."

Another serious limitation to the applicability of this research to practice is that the context in which the question about teaching have been asked must be considered. It is not simply "students" but "students in the highly structured situation of the school as it has existed in the last century" who have been queried about teaching. For example, students appear to value rather highly the ability of the teacher to control the class. In a totally different type of school situation, such as the free school, will this control still emerge as a valued teacher attribute? It might. The point is that we simply do not know.

A third limitation of this research is that much of it is either "arm-chaired" or else, at the other extreme, so inductive as to be totally unorganized. The discrepancy between the two approaches is often glaring. For example, at the lower levels of school grades, inductive remarks (where they are gathered) seem to include many references to dimensions which are absent from the usual measuring instruments in this field. For instance, in one report in which the direct interview quotes from young children are included (Leeds, 1954), of 42 negative comments made about teachers, 22 bear directly on teacher's out-of-control emotionality, especially the behavior of "hollering" and "screaming." (This is in agreement with the writer's own limited "testing" of neighborhood children, who invariably refer to "yelling and screaming.") A notable exception to this criticism is Amatora's Diagnostic Teacher Rating Scale (Amatora, 1950), but this scale is not designed for use below the fourth grade level.

It is easy to appreciate the difficulties involved in trying to correct this "preconceived versus unorganized" aspect of teacher-rating measurement. If the experimenter does not begin with some preconceived dimensions of teacher behavior, he ends with having to logically classify a great number of inductive remarks. And logical classification tends to vary from one experimenter to the next. It is for this

precise reason that it is very difficult to know how much agreement really does exist in the present literature between the varying approaches to the study of students' views of who is the "good teacher".

Reasonableness of Doing This Research

The question of whether it is even reasonable to do research in the area of young students' perceptions of teacher behavior rests in large part on whether such students can discriminate, whether their discriminations are reliable, and whether their discriminations are valid.

Can Young Students Discriminate? Much of the research in the area of teacher rating has been carried out with students at the college or high school level. At that level, data are easier to gather because verbal ability and ability to follow instructions are not problems. Furthermore, one can assume that such students are reasonably capable of making rather mature discriminations between behaviors.

Probably many teachers feel (or like to feel) that young students are not capable of making such discriminations. One example of this attitude was found in the admonition of such a teacher that "Youth should learn from the teacher and disregard his personal like or dislike of her. He should get from school

the things he has come for ...children have degrees of mentality....Their judgments of teachers are often on a level we must ignore" (Wilson, 1948, p. 65).

The evidence does not appear to support such a comforting rationalization. Of course there are areas in which the views of a particular child may not have a high correspondence to reality. In the main, though, it appears that even young students are able to agree in general about what is going on, especially when such observations are kept at the level of overt behavior.

Amatora (1952) obtained the ratings of 1,000 children in grades 4-8 on seven teachers. The students exhibited markedly different degrees of liking for three of the teachers and registered liking to almost an equal degree for the other four. Although the students were much more discriminating in their evaluation of the individual attributes of the teachers that they liked equally well, they still showed a marked ability to discriminate these same attributes in teachers whom they affectively felt quite differently about. For example, they rated a teacher whom they liked collectively at only the .02% level as almost equal in "kindness, friendliness, and understanding" with a teacher whom they liked collectively at the 65% level.

The responses of 6th graders in minority groups were not this discriminating, however, according to Pittman (1952), who

found a definite trend for evaluation to pile up in the favorable categories. Whether this is due to real lack of discrimination, reticence to express negative responses to authority figures, or even an actual high degree of favorable responsiveness to teachers is unclear. That minority groups express favorable responses to some aspects of school which more middle-class children do not view so favorably, and that they do make discriminations between various attributes of school life is demonstrated by the attitude research of Neale and Proshek (1967).

Children's discriminatory abilities are indeed (like everyone else's) affected by certain stereotypes and by stimulus generalization. Davis and Slobodian (1967), for example, could find no overt support for children's significantly expressed beliefs that boys were not allowed as much opportunity to read as girls and that boys received more negative comment from the teacher than girls. Biber and Lewis (1949), using a TAT-like testing situation, concluded that children's perceptions do actually reflect the procedures which are carried out in school, but that there is also much transference apparent in the attitudes of the children toward the teachers. In spite of this transference, however, they found that children do respond to individual differences even within a general pattern of stereotypic behavior: this was demonstrated by the children's differential responses to

one teacher in a Negro school whose behavior was atypical and did not go unnoticed as such by the children. Debus (1957) found a significant degree of transference in 5th and 6th grade students between attitudes toward their past teachers and attitudes toward their present teachers. Cox (1962) even claims a significant transference effect between attitudes toward parents and attitudes toward peers.

Wright and Sherman (1965) were interested in discovering the nature of those dimensions upon which children most agree and those areas in which they show the highest degree of disagreement. They concluded that there is very little disagreement between children in assessing a teacher's competence, but the real disagreement between children comes in their attempts to assess the teacher's affective feelings for themselves. Wright and Sherman's data can be re-analyzed, however, on the level of observation which is called for by the students. One of the present writershas done this and finds that the "competence" items are generally expressed in very concrete, observable terms, while the "love" items are generally expressed in much more removed, motivational terms. In the absence of the concreteness expressed in the former class of items, it is natural for the "love" items to lead to more idiosyncratic responses. The interpretations of both Wright and Sherman and the present writer,

however, are not necessarily at variance with the actual situation as expressed by the instrument: probably the "love" attributes of the teacher are much more covertly and individually expressed to students than the "competence" attributes of the teacher. It is important to note, however, that even in the area of highest individual disagreement --the "love" area-- the children as a class can arrive at rather reliable ratings. It therefore appears that a fruitful assessment approach in the future would consider the "love" or affective ratings of a student to be highly diagnostic when these ratings disagree with the mean values the class assigned to a teacher.

To summarize the results of studies cited above on young children's abilities to discriminate, it would seem quite sound to presume that these abilities are entirely adequate--on a group basis--for assessing the quality of teacher-pupil interaction present in a classroom. It would also appear that a valuable source of diagnostic information on any one pupil, or any subgroup of pupils, would be the degree to which these students did not agree with the rating values assigned to a teacher by the majority of students within these students' own group.

Are Ratings of Young Students Reliable? Ratings of teachers by students at the upper levels of schoolings have proven to be remarkably reliable (for example, "The Teacher's Image is

Stubbornly Stable," Bryan, 1966). Somewhat less is known about the reliability of such ratings at the lower levels of schooling. One classic attempt to resolve this question was Amatora's (1954) study of the stability of ratings of teachers done by 1,174 students in the 4th to 8th grades. Amatora found these ratings to demonstrate both a satisfactory degree of reliability and discrimination. At the present time, there is no reason to suspect that the ratings of young students are any less reliable than the ratings of older students.

Are Ratings of Young Students Valid? Do the ratings of young students correlate with other apparently valid methods for evaluating teachers? At the level of very young students, not much literature is available on this point. Teachers at the secondary level seem to feel, at least implicitly, that student ratings are a valid source of feedback, if their willingness to change on the basis of such ratings is considered. Tuckman and Oliver (1968) found that teacher behavior, which was negatively affected by feedback from supervisors, was positively improved by student feedback. Teachers seemed to acknowledge the validity of student feedback by their differential reaction to it.

Perhaps the research on the validity of young students' ratings of peers can be accepted, at least partially, in support

of the general validity of young person's perceptions. Quite a bit of evidence (Bonney, 1955; Cannon, 1958; Fine, Fulkerson, and Phillips, 1955; Gronlund and Whitney, 1956; Kuhlen and Collister, 1952; and Phillips and DeVault, 1955) exists to demonstrate the high validity of children's perceptions as a criterion, taken collectively, of other children's behavior. In fact Bower (1969) states, "If only one method for class analysis were permissible, this (a sociometric instrument completed by children) would undoubtedly be the best single procedure (p. 72)". At least for the reasons stated by Kratz in 1896--"...those characteristics which impress the pupils favorably, which lead to a high appreciation on their part, and which establish those relations of sympathy and cooperation so essential in the schoolroom, must have some value...(p. 413)"--there appears to be no real question of the validity or usefulness of young children's perceptions of the teacher.

Methodology

Having established at least tentatively the reasonableness of assessing teacher behavior by obtaining young students' perceptions of it, a next important concern is the discovery of suitable methodology for making these assessments.

At the upper educational levels, it is usual to administer to the student some type of verbal scale which describes certain

attributes of teacher behavior and asks for a student's reactions in terms of magnitude of his own feelings (for example, the SET, Veldman, 1970). Obviously, such an approach presents difficulties with very young students who often cannot read.

Table 1 summarizes a representative sample of the major methodological approaches which have been used in securing student responses to teacher evaluation below the upper high school level. It is apparent that scales are rarely used below the fourth grade. Written responses are secured below this level only with gifted students (Barbe and Steiert, 1964). The major approaches which have been used with students below the fourth grade are TAT-like instruments and interviews recorded by an interviewer. Drawings and the Semantic Differential have been used infrequently. (Please see Table 1 for references.) This implies that what is known, in an organized fashion, about teacher evaluation by students at the lower levels of schooling is primarily that information gleaned from the fourth to eighth grades. Views of what attributes constitute "the good teacher" at the very lowest level of schooling, i.e., the kindergarten and first three grades, are in a much more inductive and unquantified condition.

TABLE 1

Methods of Eliciting Student Ratings of Teachers

(Below Upper High School Level)

	Grade Level									
	K	1	2	3	4	5	6	7	8	9
I. INDUCTIVE VERBAL RESPONSES:										
Interviews, remarks recorded by interviewer:										
Davis & Slobodian, 1967	X									
Spillard, 1964	X	_____	_____	_____	_____	_____	_____	_____	_____	X
Remarks, student written:										
Leeds, 1954						X	___	___	___	X
Barbe & Steiert, 1964 (Gifted S's)	X	_____	_____	_____	_____	_____	_____	_____	_____	X
Essays:										
Taylor, 1962	"Grammar school"	_____	_____	_____	_____	_____	_____	_____	_____	"Secondary"
II. DRAWING:										
Gregerson & Travers, 1968	X	_____	_____	_____	_____	_____	_____	_____	_____	X
III. TAT-LIKE INSTRUMENTS:										
	"Young children"						"Pre-adolescents"			
Cohen, 1967										
Cox, 1962							X	__		X
Biber & Lewis, 1949	X	_____	_____	_____	_____	_____	_____	_____	_____	X
Bower, 1969										

TABLE 1 (cont.)

	Grade Level K 1 2 3 4 5 6 7 8 9 10 11 12
IV. SEMANTIC DIFFERENTIAL:	
DiVesta & Dick, 1966 (reliability study)	X_____?
Neale & Proshek, 1967	X____X
V. SCALES:	
Diagnostic Teacher Rating Scale:	
Amatora, 1950, 1952, 1954 (3 studies)	X_____X
Burkard, 1965	X_____X
Pupil Observation Survey:	
White & Dekle, 1966	X____X
Veldman & Peck, 1963	X_____X
Rogge's Scale of Teaching:	
Paraskevopoulos, 1968 (Gifted S's)	"Elementary"_____"Secondary"
Illinois Inventory of Pupil Opinion:	
Pittman, 1962 (Negro S's)	X

TABLE 1 (cont.)

	Grade Level								
	K	1	2	3	4	5	6	7	8	9
Various Author-Originated Scales:										
Christensen, 1960				X		X				
Gage, et. al., 1955				X				X		
Wright & Sherman, 1965				X						X
Tenenbaum, 1944 (Includes lower SES S's)							X	X		
Reed, 1961									X	
VI. TRUE-FALSE FORMAT:										
Ojemann & Snider, 1963				X						X

Results

Some of the major goals of persons doing research within the area of student perceptions of teachers have been these: (1) the study of the general positiveness or negativeness of student attitudes towards teachers and schools, (2) the search for basic dimensions of teacher behavior which seem to determine the positive or negative quality of the student's attitude toward the teacher, (3) certain teacher attitudes which correlate with student attitudes toward teachers, and (4) certain student outcomes, such as achievement, which correlate with either the students' views of the teacher or the teacher's attitudes.

The results reported here will concentrate on No. 2 above, i.e., the search for basic dimensions of teacher behavior which seem to determine the positive or negative quality of students' attitudes toward teachers. These are the behaviors which will need to be tapped in a new instrument. A few salient teacher attitudes (No. 3) related to student opinion will be examined also. As with the rest of this review, the discussion will be confined mainly to results obtained in studies done with students below the high school level.

Basic Dimensions of Teacher Behavior. The search for major dimensions of teacher behavior which appear to determine the positiveness or negativeness of students' attitudes is usually

conducted via the methodology of factor analysis, as we shall see below:

Leading efforts to identify important factors in teacher behavior are those of Ryans (1960), who identified three major factors, and Veldman and Peck (1963), who developed a Pupil Observation Survey yielding five major factors. The factors of Ryans were these: (1) Pattern Xo: understanding, friendly versus aloof, egocentric, restricted; (2) Pattern Yo: responsible, businesslike, systematic versus evading, unplanned, or slipshod; (3) Pattern Zo: stimulating, imaginative, surgent or enthusiastic versus dull, routine. The five factors in teacher behavior identified by Veldman and Peck were these: (1) friendly, cheerful, admired; (2) poised, knowledgeable; (3) interesting, preferred; (4) strict control; and (5) democratic procedure. Veldman and Peck's subjects were junior and senior high school students.

When Veldman and Peck's Pupil Observation Survey (POSR) was used by White and Dekle (1966) at the 5th to 7th grade level, six factors were extracted which these authors called (1) warm, affable, deferring; (2) fair, considerate; (3) controlled, orderly; (4) surgent, stimulating; (5) knowledgeable, open-minded; and (6) self-assured, poised, ego-strength. White and Dekle further divided their 158 Ss into three groups of over-, under-, and normally-achieving students. They found a wide difference

between over- and under-achieving students' behavior on Factor 1 (and Factor 1 only), which was the students' perception of the warmth and affableness of the teacher.

Another attempt to extract the basic dimensions of teacher behavior from factor analysis of student ratings was that of Paraskevopoulos (1968), who studied teacher ratings in a sample of gifted children. Three factors emerged from this author's factor analysis of Rogge's Style of Teaching Inventory: (1) friendly, warm; (2) understanding, flexible; and (3) encouraging initative and participation. Paraskevopoulos further discovered that these factors related differentially to other teacher characteristics such as measured by the Myers-Briggs Inventory, for example, sensing, intuiting, feeling, and thinking propensities. These results would lead one not to expect all the positive "factors" to be present in one teacher.

Neale and Proshek (1967) used a Semantic Differential in a factor analytic approach with 4th to 6th graders. They also discovered three factors, but the first--evaluation--was by far the most important one. Their study is important because a part of it was done with a culturally disadvantaged elementary school population, and they did not find a generally negative across-the-board attitude toward teachers or school among these children.

Still another factor-analysis approach to the study of basic teacher dimensions as seen by students was done by Beck (1967) using his own questionnaire "About My Teacher" with 2,108 sixth grade students. He extracted one strong major factor in teacher behavior which, upon analysis, was composed of the following elements: (1) warmth and friendliness, (2) ability to communicate clearly, and (3) motivating qualities. Though Beck's second and third factors were statistically less important, they appeared to be two separate dimensions of student perception about discipline: One concerned discipline as reflecting behavior of students, and the other concerned discipline as originating from the teacher. Beck considers this result of possible importance for future research, since it appears to indicate that students can differentiate between situational aspects of disciplinary problems.

A somewhat different approach from factor analysis was taken by Amatora (1950), who began the search for major dimensions of teacher behavior with an inductive study that served as the basis for the development of a quantified measuring scale. As cited above, Amatora's approach is closer to the ideal than most. First, 200 elementary school students were asked to list traits which they liked and disliked in their teachers. A list of 1,500 items resulted, from which 300 items were edited into the following seven scales: (1) pupil liking for the teacher, (2) teacher's

ability to explain, (3) the sympathetic nature of the teacher, (4) fairness, especially in grading, (5) discipline, (6) amount of work required, and (7) pupil liking for lessons taught by the teacher. The final form of these scales was named the Diagnostic Teacher Rating Scale, with produced reliabilities between .72 and .95 in a sample of 1,174 students, and low correlations between the separate scales in a sub-sample of 300 students. Further research with the instrument has since been done in grades 4 and above (see Table 1).

An example of a purely inductive and logical approach to the search for dimensions of teacher behavior is that done by Barbe and Steiert with gifted children in grades 3 - 6 (1964). The concern these particular children listed most often (36 out of 40 Ss) related to the knowledge and wisdom possessed by the teacher. Next in importance (21 out of 40) was the children's desire for firm class control tempered usually with "fairness" or "kindness". These boys wanted a considerate, older woman teacher; the girls wanted a teacher who would recognize their limits. In short, these gifted children asked for knowledgeable, warm teachers--but they were unique of all the studies surveyed in putting the dimension of teacher knowledge first.

In a different approach to the isolation of important teacher characteristics, Symonds (1955) combined a student-nomination

technique with his own subjective observations of the classroom behavior of 17 teachers taken from the top and bottom of the student's nomination list. He concluded as follows: (1) the superior teachers liked children, and the inferior teachers disliked children; (2) the superior teachers were personally secure and self-assured; the inferior teachers were insecure and had feelings of inferiority and inadequacy; (3) the superior teachers were well-integrated and possessed good personality organization; the inferior teachers were personally disorganized. He included anecdotal descriptions for all six of his conditions.

In the above studies which have been cited as representative examples of the quest for basic dimensions of teacher behavior that appear to affect children's attitudes positively or negatively, it is rather difficult to determine (as has been previously discussed in this paper) exactly where agreement and disagreement exists. About the clearest result is this: almost all such attempts find a common result in the primary appearance of a teacher behavior usually referred to as "warmth and friendliness". In factor analytic studies, this is the usual meaning of the first factor extracted.

After the identification of this first factor, results vary. (For a visual comparison of some of those results, see Table 2.) The practice of naming factors may, or may not, account for

TABLE 2

Comparison of Factors Identified

Veldman and Peck, 1963	Friendly Cheerful	Knowledge Poised	Lively Interesting	Control	Democratic					
Amatora, 1950	X		(Liking for lessons)	X		Can Explain	Sympa-thetic	Fair-ness	Amt. Work	
White and Dekle, 1966	X	X/X (Separate)	X	X				X		
Barbe and Steiert, 1964		X						X		
Beck, 1967	X					(Communi-cate)				Moti-vate

Key:

X indicates the same factor as named immediately above

Factor names in parentheses appear under factors with which they probably agree

variability in results. For example, it is hard to tell whether one experimenter's factor of "ability to motivate" really is the same dimension as the next experimenter's "lively and interesting". Logically, they appear to be quite the same type of behavior.

It is a further complication for one interested in knowing the teacher characteristics which are appreciated by younger children to note that only one of the studies done with younger children (below 4th grade) appears in any of the "major dimensions" results cited above. Most of the experimenters who studied young children's attitudes toward teachers were not primarily interested in the issue of the "major dimensions" of teaching. For example, Gregerson and Travers' study (1968) using a drawing technique was concerned mostly with sex differences in attitudes, Davis and Slobodian's study (1967) was concerned with the accuracy of children's perceptions about discrimination against boys, Biber and Lewis' study using a TAT-type testing situation was primarily concerned with transference of attitudes in a school of lower socioeconomic level, and so forth. Spillard (1964) did logically organize the inductive remarks she secured from young students into these exhortations to teachers: (1) smile a bit, (2) don't be a lemon, (3) your key word should be justice, (4) social attitudes are contagious, and (5) we need effective teaching methods. Reanalysis of Spillard's groups of quotations

leads one to believe her first grouping deals with the commonly-found dimension of warmth and friendliness; the second is a dislike of teacher anger factor; the third grouping is a demand for fair but consistent control; the fourth has to do with a desire for getting one's share of personal interaction with the teacher (at the level of getting to run errands, etc.)--a kind of "attentional" factor which seems appropriate for such young children with the teacher largely viewed as a mother figure; and the fifth is a rather perceptive demand for more "innovation": i.e., opportunity to do things on one's own, move around, and in general, take a larger part in the learning activities. Spillard's N was small (20), but her use of direct quotations makes the study provocative.

In conclusion, at this stage in the research on teacher characteristics from the pupils' point of view, the one statement that can be made with a high degree of confidence is that children of all ages want a warm and friendly teacher who likes kids. Nowhere is there evidence to suggest that children do not also desire teacher competence, however. Children themselves apparently see no contradiction between a strong commitment to learning and good mental health conditions in the school (Dunn, Bloom, and Morse, 1963). They give indications in almost all research cited, though it is not stated with striking verbal

agreement from one study to the next, that they want the teacher to know something and be able to communicate this knowledge in a way which involves and interests them. The following, while an over-simplication, is a principle which appears supported by the literature: the necessary condition of good teaching appears to be warmth and friendliness toward children, while the sufficient condition of good teaching appears to be competence.

A Major Teacher Attitude Gestalt Which Appears to Relate to Student-Identified "Warmth and Friendliness"

Though the major concern of this paper is the investigation of teacher behaviors rather than teacher attitudes and their correlates with student opinion, one attitude gestalt identified in teachers has appeared in the literature investigated with such force that tracking it down may explicate those teacher behaviors which pupils regard as "warm and friendly". This is the teacher attitude gestalt of ego-centricity.

At the conclusion of Ryans' book on teacher characteristics (1960), he makes the following summary remarks about good and bad teachers:

> There was a general tendency for high teachers to be extremely generous in appraisals of the behavior and motives of other persons...On the other hand, low teachers tended generally to be restrictive and critical in their appraisals of other persons....(pp. 397-398)

Burkard (1962) found very significant differences in the TAT sequence analysis of the fantasy productions of teachers rated high and low by students using the Diagnostic Teacher-Rating Scale. She states, "...there is a strong contrast between what may be designated the active attitude of the high group and the passive attitude of the low ".

> The teachers in the high group constantly reiterate the need of action on their part to achieve success or to overcome failure. They foresee difficulties and recognize the need of planning and preparation. For them success follows on such activity. The low group, on the other hand, do not mention these requirements of success. According to their stories, success just comes easily or in some equally unrealistic way; or, again it does not come at all in spite of constructive activity on the achiever's part. . .
>
> Toward other people, the high-ranked group expresses a willingness to go along reasonably, seeing them as good and helpful. The low group, on the contrary, sees people either as hostile or as friendly in a sentimental way. (p. 285).

Burkard (1965) followed this study with another investigation of teacher attitudes, again using student ratings as the criterion. The following were items on the Minnesota Teacher Attitude Inventory upon which the groups of high and low teachers differed significantly (with high teachers on the positive end): teaching never gets monotonous; most pupils are considerate of their teachers; to maintain good discipline in a classroom a teacher does (not) need to be "hardboiled"; the majority of children take their responsibilities seriously; most children

(are courteous) to adults; young people of today are just as good as those of past generations; and most pupils try to make things easier for their teacher. These are obviously generous, other-centered statements.

Finally, Symonds (1955), as previously cited, found good and poor teachers to differ markedly in liking for students and in ego-functioning terms.

The studies cited above contain a remarkable degree of agreement when viewed conceptually from the framework of ego psychology. The warm and friendly teacher who is rated as superior by the students is a mature adult whose focus is outwardly directed toward the children and, furthermore, a person who views the children in a very positive and generous kind of light. The poor teacher appears to be thoroughly ego-centric, concerned with herself, interpreting the students' actions as personally directed toward her own discomfort, and disposed to impugn the motives of others.

These studies appear important because they help to delimit the rather broad conceptual reaction of students, i.e., the teacher is "warm and friendly," to more definable teacher attitudes which precipitate these emotional reactions in students.

CHAPTER III

DEVELOPMENT OF THE STUDENT EVALUATION OF TEACHER (SET II)

Evidence has been presented on the usual methodology for securing teacher ratings at the lower grade levels in school. (See Table 1 in the preceding Review of the Literature section.) It is apparent that scaling techniques have not often been employed below fourth grade, and that the usual procedures for teacher rating below this grade have involved one-to-one testing.

A new teacher rating instrument--the SET II--was designed by one of the present writers (Dr. Ruth A. Haak) to assess children's views of teachers below fourth grade without the necessity for individual testing. The items for this instrument were modeled after items in the "My Teacher" scale developed by Wright and Sherman (1965) from Redl's leader-tyrant typology. The reasons for selecting items like those of this particular measure were these:

(1) Wright and Sherman had identified two major dimensions elicited by their instrument when used with elementary school children: the "love" dimension and the "competency" dimension. From the research review presented, these two dimensions appeared to be, at least, the major ones underlying children's judgments of teachers.

(2) This instrument might be able to yield both reliable mean ratings to be used as criteria of teachers for research

purposes and diagnostic information for use in working with individual children in the school setting. This seemed possible since Wright and Sherman had not only identified two major dimensions which children rate reliably--love and competency--but they also discovered that responses to some items vary sufficiently to yield valuable diagnostic information in the case of children who give deviant responses.

With considerable rewording of some of the items used by Wright and Sherman and inclusion of new value dimensions which the literature suggested were particularily important at lower grades--especially, dislike of angry responses and chance to interact with teacher--the proposed items for the new children's teacher rating scale (SET II) were as follows: (the items are presented in their hypothesized "groupings," but items were randomized before presentation).

Assessment of Teacher

Competency Items:

1. She teaches us a lot.
2. She is a smart teacher.
3. We can tell how she wants things done.

Love Items:

4. The kids like her.
5. She makes school fun.
6. She always picks on people.

Diagnostic (?) Items:

7. She helps us a lot.
8. She gets mad a lot.
9. She listens to what we want.
10. She gives us too much work.

Finally, an innovation in teacher-rating instrumentation suggested by Don Veldman (The Set I., Veldman, 1970), was included, namely a second set of items analogous to those already presented which attempted to tap the child's belief about his teacher's supposed attitudes toward himself. The dimensions for this second section were similar to those used to rate the teacher. This second section of the SET II was admittedly partly "projective," of course, but seemed of especial importance in view of the emphasis that the literature review had suggested for the teacher's ability to communicate a "generous view" of the student to the student himself (see the section on Egocentricity, Chapter 2, this manual).

This second section of the SET II contained the following items:

Estimation of Teacher's Attitude

Competency Items:

12. She thinks I work hard.
13. She thinks I am smart.
*22. She thinks we are a smart class.
14. She thinks I can do a lot on my own.

*Item 22 is a group form of Item 13.

Love Items:

15. She likes me.
*23. She likes us kids.
11. She likes to teach.
17. She thinks that kids are good.

Diagnostic (?) Items:

18. She likes for me to help her.
19. She thinks I act ugly.
20. She thinks I have good ideas.
21. She thinks I am lazy.

* Item 23 is a group form of Item 15.

The total of twenty-two items were printed upon small cards with an identifying "stamp" on the upper right hand corner of each card. (See Appendix A for an illustration.) The first set of items (Items 1-10) was printed on yellow paper, the second set (Items 12-23) on blue paper.

When the test is administered, the tester orally identifies each card by its "stamp" to the children. The wording of the items is printed upon the cards merely for its face validity value. (Some children can read the wording, of course, but the ability to do so is not necessary). The tester then reads the item aloud, and the child classifies the item on each card as being true or false by placing the card in one side of a two-sided sorting envelope. On one side of the sorting envelope appears the picture of a post-office box: this is where

the card is to be placed if the child thinks that what the card "says" is true. On the other side of the sorting envelope appears a picture of a wastebasket: this is where the card is to be placed if what the card "says" is not true. (See Appendix B for a picture of the sorting envelope.) An entire class of children can be tested at one time by one test administrator. The average time involved is about 20 minutes at kindergarten and first grade level and approximately 10 minutes at grade levels above first.

The first large scale testing for this instrument was conducted in the Austin, Texas public schools. Approximately 1,040 children were involved in a pre-post reliability design. This initial testing is reported in the next section.

CHAPTER IV

THE STUDENT EVALUATION OF TEACHER II (SET II)

In April, 1971, 1,040 children in grades K-6 in the Austin, Texas public schools served as subjects for the initial pilot study of the SET II. This included a test of the instrument's stability.

All children were tested in a classroom situation by a single test administrator with no other adult (nor the teacher) present in the room. The time lapse between testing and re-testing was 10 days. In nearly all cases, the testing in both administrations was done by the same person. A total of four test administrators were involved in the entire testing operation.

The sample of children involved in this study was pre-planned to include a stratified representation of school grades and cultural groups. In Austin, Texas, there are three major cultural groups: Black, Mexican-American, and White. (Racially, Mexican-Americans are classified as "white"; however, important cultural differences exist between Mexican-Americans and other white groups. In these data, therefore, "white" is understood to be a classification for all whites other than Mexican-American.)

Reliability Study

Since the SET II is presented in a true-false format, the method selected for determining reliability with these data was a percentage of agreement statistic. The percentage of agreement (PA) reliability figures for all children for all items are presented in the following formats: PA by grade only (See Table 3); PA by grade with cultural group and ability level subgroupings (See Tables 4-6); and PA by ability level within cultural group (See Tables 7-8).[1] All reliabilities for all items were considered to be adequate at this time to justify their continuance in the experimental SET II except for the reliabilities obtained at certain ability levels within the kindergarten sample. Because of the difficulties with the kindergarten sample in these data, further analyses of the SET II did not include the data from kindergarten children. Use of the SET II with kindergarten children would presently be recommended only under very limited conditions--specifically, with children who are average or above in their classroom ability and performance. This instrument in its present form does not appear satisfactory for use with lower

[1]Caution should be used in interpreting the ability-within-cultural group PA's. These data are not from a sufficiently large N to justify their acceptance and are presented mostly for the hypotheses they may generate.

TABLE 3

Percentage of Agreement Pre to Post:
Total Sample at Each Grade Level

Children in Kindergarten		Children in First Grade	
1	86.0	1	89.6
2	84.9	2	92.2
3	69.2	3	70.1
4	89.5	4	88.7
5	74.6	5	83.5
6	63.0	6	78.4
7	78.0	7	86.6
8	59.3	8	69.7
9	65.9	9	81.7
10	63.4	10	67.4
11	74.6	11	86.6
12	74.6	12	89.1
13	73.9	13	91.3
14	69.9	14	80.5
15	81.9	15	86.6
16	87.3	16	87.0
17	82.7	17	89.6
18	74.6	18	80.1
19	78.0	19	86.6
20	67.6	20	73.2
21	68.2	21	79.8
22	67.1	22	75.3
N =	172	N =	231

TABLE 3 (cont.)

Second and Third Grades		Fourth-Sixth Grades (Mexican American)	
1	88.4	1	79.2
2	88.4	2	83.6
3	76.1	3	75.3
4	82.7	4	84.9
5	83.4	5	83.6
6	79.7	6	80.8
7	89.4	7	81.5
8	68.4	8	66.9
9	74.4	9	78.9
10	68.3	10	77.2
11	77.2	11	69.1
12	83.7	12	77.5
13	86.1	13	82.7
14	79.1	14	76.9
15	86.8	15	85.6
16	88.1	16	90.8
17	86.1	17	94.0
18	82.1	18	84.9
19	78.8	19	78.8
20	72.8	20	76.4
21	74.8	21	69.3
22	72.8	22	83.2
N =	301	N =	152

TABLE 4

Children in Kindergarten
Percentage of Agreement Pre to Post

By Cultural Group: Percent Agreement T-1 to T-2				By Ability Level:* Percent Agreement T-1 to T-2			
Variable	Black	Mexican-American	White	Variable	Low Ability	Medium Ability	High Ability
1	69.0	87.5	93.3	1	80.0	86.2	89.6
2	78.6	76.9	91.2	2	77.5	89.1	85.3
3	59.5	66.7	74.7	3	67.5	71.9	67.6
4	80.9	95.0	91.1	4	87.5	89.2	91.0
5	66.7	67.5	81.3	5	72.5	76.9	73.5
6	52.4	77.5	61.5	6	57.5	56.9	72.1
7	57.1	87.5	83.5	7	77.5	75.4	80.9
8	57.1	37.5	70.0	8	62.5	56.9	59.7
9	57.1	55.0	74.7	9	60.0	70.8	64.7
10	69.0	55.0	64.4	10	65.0	60.0	65.7
11	69.0	87.5	71.4	11	85.0	75.4	67.6
12	61.9	72.5	81.3	12	80.0	69.2	76.5
13	54.8	80.0	80.2	13	82.5	69.2	73.5
14	66.7	62.5	74.7	14	70.0	69.2	70.6
15	73.8	74.4	89.0	15	95.0	78.5	77.6
16	78.6	85.0	92.3	16	87.5	84.6	89.7
17	66.7	82.5	90.1	17	92.5	80.0	79.4
18	66.7	55.0	86.8	18	82.5	73.8	70.6
19	61.9	80.0	84.6	19	90.0	73.8	75.0
20	52.4	57.5	79.1	20	77.5	61.5	67.6
21	64.3	57.5	74.7	21	72.5	66.2	67.6
22	64.3	65.0	69.2	22	65.0	70.8	64.7
N =	42	40	91	N =	40	65	68

*Ability level is measured by teacher's estimation of the class.

TABLE 5

Children in Grades 1-3
Percentage of Agreement Pre to Post

By Cultural Group:				By Ability Level:			
Percent Agreement T-1 to T-2				Percent Agreement T-1 to T-2			
Variable	Black	Mexican-American	White	Variable	Low Ability	Medium Ability	High Ability
1	94.6	83.3	90.6	1	85.4	88.2	92.6
2	87.8	90.1	91.7	2	88.5	88.8	92.6
3	75.7	68.5	77.3	3	66.9	74.9	77.7
4	85.1	82.3	88.9	4	84.7	81.8	89.4
5	80.9	83.7	85.1	5	79.5	79.1	90.9
6	78.4	75.4	83.9	6	77.1	74.3	85.6
7	85.8	90.6	87.3	7	84.1	86.6	93.1
8	60.3	67.2	76.8	8	62.6	64.9	77.1
9	80.9	75.9	76.8	9	73.7	74.3	84.0
10	63.0	64.5	75.7	10	60.9	68.8	72.9
11	85.8	78.4	80.7	11	74.5	81.9	86.2
12	93.2	79.8	87.3	12	82.1	85.6	89.9
13	90.5	86.3	88.9	13	86.6	87.8	90.4
14	83.8	72.1	85.1	14	73.9	76.6	87.8
15	90.5	80.9	90.1	15	82.8	85.1	91.5
16	87.8	86.3	88.9	16	85.9	86.2	90.4
17	91.2	83.3	89.5	17	85.4	88.3	88.8
18	79.1	76.8	87.8	18	80.9	78.2	84.5
19	85.8	78.9	82.9	19	77.1	82.9	85.6
20	70.9	66.2	82.3	20	68.8	70.2	79.3
21	76.7	70.6	84.4	21	75.6	73.3	81.7
22	72.1	65.2	85.1	22	68.2	72.3	80.2
N =	148	204	181	N =	157	188	188

TABLE 6

Children in Grades 4-6*
Percentage of Agreement Pre to Post

Variable	By Ability: Low Ability	Medium Ability	High Ability
1	80.0	76.7	81.6
2	77.3	85.0	87.5
3	67.4	76.3	81.3
4	82.2	89.7	81.6
5	79.1	85.0	85.7
6	79.1	77.9	85.7
7	74.4	81.4	87.8
8	62.2	63.3	75.5
9	72.7	77.9	85.7
10	65.9	80.0	83.3
11	59.1	68.3	79.2
12	71.1	71.9	89.8
13	79.1	84.5	83.7
14	72.1	71.2	87.8
15	86.7	79.7	91.8
16	86.4	95.0	89.8
17	93.0	91.5	98.0
18	84.1	83.1	87.8
19	79.5	77.6	79.6
20	72.1	69.0	89.4
21	56.8	71.7	77.6
22	86.0	73.7	91.8
N =	45	60	49

*All children in the Grades 4-6 sample are Mexican-American.

TABLE 7

Children in Kindergarten
Ability Level Within Race
Percentage of Agreement Pre to Post

Black				Mexican-American			
Variable	Low Ability	Medium Ability	High Ability	Variable	Low Ability	Medium Ability	High Ability
1	70.0	61.1	78.6	1	85.7	88.2	87.5
2	70.0	77.8	85.7	2	71.4	81.3	75.0
3	50.0	66.7	57.1	3	57.1	56.3	81.3
4	70.0	83.3	85.7	4	100.0	88.2	100.0
5	60.0	61.1	78.6	5	71.4	70.6	62.5
6	50.0	44.4	64.3	6	57.1	82.4	81.3
7	60.0	44.4	71.4	7	85.7	76.5	100.0
8	50.0	61.1	57.1	8	28.6	41.2	37.5
9	40.0	61.1	64.3	9	71.4	70.6	31.3
10	60.0	55.6	92.9	10	28.6	64.7	56.3
11	80.0	55.6	78.6	11	85.7	88.2	87.5
12	70.0	55.6	64.3	12	71.4	52.9	93.8
13	70.0	38.9	64.3	13	100.0	64.7	87.5
14	60.0	61.1	78.6	14	71.4	64.7	56.3
15	100.0	55.6	78.6	15	85.7	76.5	66.7
16	80.0	72.2	85.7	16	71.4	76.5	100.0
17	80.0	61.1	64.3	17	100.0	88.2	68.8
18	80.0	44.4	85.7	18	42.9	76.5	37.5
19	100.0	38.9	64.3	19	85.7	82.4	75.0
20	60.0	38.9	64.3	20	71.4	64.7	43.8
21	70.0	55.6	71.4	21	71.4	52.9	56.3
22	60.0	61.1	71.4	22	57.1	70.6	62.5
N =	10	18	14	N	7	17	16

TABLE 7 (cont.)

White

Variable	Low Ability	Medium Ability	High Ability
1	82.6	100.0	94.6
2	82.6	100.0	89.5
3	78.3	83.3	65.8
4	91.3	93.3	89.2
5	78.3	90.0	76.3
6	60.9	50.0	71.1
7	82.6	93.3	76.3
8	78.3	63.3	70.3
9	65.2	76.7	78.9
10	78.3	60.0	59.5
11	86.9	80.0	55.3
12	86.9	86.7	73.7
13	82.6	90.0	71.1
14	73.9	76.7	73.7
15	95.7	93.3	81.6
16	95.7	96.7	86.8
17	95.7	86.7	89.5
18	95.7	90.0	78.9
19	86.9	90.0	78.9
20	86.9	73.3	78.9
21	73.9	80.0	71.1
22	69.6	76.7	63.2
N =	23	30	38

TABLE 8

Children in Grades 1-3
Ability Level Within Race
Percentage of Agreement Pre to Post

Black				Mexican-American			
Percent Agreement T-1 to T-2				Percent Agreement T-1 to T-2			
Variable	Low Ability	Medium Ability	High Ability	Variable	Low Ability	Medium Ability	High Ability
1	88.9	94.4	100.0	1	80.0	84.5	84.7
2	93.3	81.5	89.8	2	85.0	91.5	93.1
3	73.3	75.9	77.6	3	58.3	70.4	75.0
4	86.7	77.8	91.8	4	81.7	84.5	80.6
5	75.0	74.1	93.9	5	80.0	78.9	91.7
6	73.3	70.4	91.8	6	78.3	73.2	75.0
7	82.2	79.6	95.9	7	86.7	91.5	93.1
8	53.5	53.7	73.5	8	61.7	66.7	72.2
9	79.5	72.2	91.8	9	78.3	71.8	77.8
10	54.5	67.9	65.3	10	55.0	70.4	66.7
11	84.4	87.0	85.7	11	66.7	81.9	84.7
12	86.7	96.3	95.9	12	75.0	78.9	84.7
13	88.9	92.6	89.8	13	86.7	83.3	88.9
14	77.8	81.5	91.8	14	63.3	69.4	81.9
15	82.2	92.6	95.9	15	75.0	80.6	86.1
16	84.4	88.9	89.8	16	88.3	83.3	87.5
17	95.6	90.7	87.8	17	78.3	84.7	86.1
18	75.6	75.9	85.7	18	78.3	73.6	78.9
19	84.4	90.7	81.6	19	65.0	83.3	86.1
20	75.6	62.9	75.5	20	58.3	63.9	75.0
21	79.5	72.2	79.2	21	66.7	66.7	77.8
22	73.3	70.4	72.9	22	55.0	66.7	72.2
N =	45	54	49	N =	60	72	72

TABLE 8 (cont.)

White

Percent Agreement T-1 to T-2

Variable	Low Ability	Medium Ability	High Ability
1	88.5	87.1	95.5
2	88.5	91.9	94.0
3	71.2	79.0	80.6
4	86.5	82.3	97.0
5	82.7	83.9	88.1
6	78.8	79.0	92.5
7	82.7	87.1	91.0
8	71.2	72.6	85.1
9	63.5	79.0	85.1
10	73.1	67.7	85.1
11	75.0	77.4	88.1
12	86.5	83.9	91.0
13	84.6	88.7	92.5
14	82.7	80.6	91.0
15	92.3	83.9	94.0
16	84.6	87.1	94.0
17	84.6	90.3	92.5
18	88.5	85.5	89.6
19	84.6	75.8	88.1
20	75.0	83.9	86.6
21	82.7	81.9	87.9
22	78.8	80.6	94.1
N =	52	62	67

ability kindergarten children. Other differences in the percentage-of-agreement reliabilities between cultural groups and grade levels did occur; however, these differences, upon examination are generally not marked; the first grade data did not suggest any reason to treat this group separately from the rest of the sample, who were older. It appears that all children in grades 1-3[2] regardless of ability level or cultural group can use the SET II to rate teachers in a reliable manner.

Factor Analyses

A large number of factor analytic studies were performed upon the SET II pilot data (See Appendix C). The procedure used provided principal-axis structures and a simple structure resulting from rotation (Veldman, 1967) The latter structures were examined with the purpose of identifying meaningful dimensions of responses.

Though some suggestive differences between cultural groups and grade levels did appear, these differences tended not to be marked and to disappear with age. In fact, these differences were not sufficient to distort a preliminary factor analysis

[2]The SET II is designed for use in grades 1-3 or 4, since the SET I can be used with grades higher than this. Nevertheless, a large sample of older Mexican American children was available in one site, and they were included in this study for purposes of testing the feasibility of the use of the SET II with a population which contained many poor readers.

performed upon all these data grouped together, regardless of grade (1-3) or cultural group. The differences which did appear were chiefly as follows:

1. The black children were much more cohesive in their evaluations of the teachers' supposed attitudes toward themselves than they were in their own evaluations of the teacher. This finding is quite provocative and suggests a number of hypotheses about the way in which black children orient their perceptions.
2. Mexican-American children tended to be more cohesive in their evaluations of items that were emotionally loaded. By late elementary school, however, this difference tended to disappear and the Mexican-American and other white data were largely indiscernible in this regard.

The major source of error to be contended with in the factor analytic studies was the high possibility that individual "teacher effects" could be strongly and undesirably affecting the factor structure. The final factor studies (reported in Tables 9 & 11) were, therefore, performed upon teacher mean scores for all variables. There were a total of 28 teachers of students in grades 1-3 involved in the study; therefore, data analyzed in the major factor structure reported in Table 9 are comprised of the mean scores for 20 of the original 22 items for each of

TABLE 9

Factor Analysis of SET II
Grades 1-3

66.7 PCT. of the variance was extracted by 3 Roots as follows:

Factor 1	Factor 2	Factor 3
30.8	18.7	17.2

Factor Loadings After Rotation to Simple Structure

Variable (Item)	Factor 1	Factor 2	Factor 3
3	.65	-.51	-.06
4	.85	-.17	-.22
5	.94	-.10	-.13
6	-.55	.64	.08
7	.70	-.25	-.18
8	-.16	.80	-.11
9	.66	-.28	-.27
10	-.11	.71	.10
11	.20	.08	-.63
12	.48	-.29	-.72
13	.71	-.16	-.30
14	.27	.05	-.71
15	.12	-.20	-.80
16	.81	-.17	-.36
17	.60	.13	-.43
18	.64	-.45	-.19
19	.15	-.45	-.59
20	-.31	.73	.19
21	.67	-.10	-.59
22	-.00	.77	.13

teachers in grades 1-3.[3] The factor study in Table 11 included all 36 teachers in the sample, those who taught grades 4-6 as well as grades 1-3. When the SET II was originally conceived, it was hypothesized that responses to the items would statistically group themselves into the following categories: Love, Competence, and Diagnostic (items with little logical commonality other than their emotional tone, and with high standard deviations). The major factor structure which resulted from the final analysis of the Experimental SET II (teacher mean scores on all variables) only partially supported these preliminary hypotheses. A discussion of the final factor structure accepted for scoring the SET II by categorical dimensions follows.

The Categorical Structure of the SET II

In this study, there appear to be three relatively stable factors upon which children in grades 1-3 make their evaluations of teachers. (See Table 10) The largest of these factors is the one to be labeled "Stimulating Interactive Style". This quality involves both child-teacher rapport and teacher competence, for young children are apparently only in the process of differentiating this factor into two components. If a four-factor solution is envoked for the data from grades 1-3, "competence

[3]Original items #1 and #2 were dropped from this final analysis for reasons discussed on page 63f.

TABLE 10

Primary (Highest) Loadings
SET II Items--Three Factor Structure
Grades 1-3

Factor	Item No.	Statement	Loading
Factor 1: Stimulating, Interactive Style	5	She makes school fun.	.94
	4	The kids like her.	.85
	16	She likes us kids.	.81
	13	She thinks we are a smart class.	.71
	7	She helps us a lot.	.70
	21	She thinks I have good ideas.	.67
	9	She listens to what we want.	.66
	3	We can tell how she wants things done.	.65
	18	She thinks that kids are good.	.65
	17	She likes to teach.	.60
Factor 2: Unreasonable Negativity	8	She gets mad a lot.	.81
	22	She thinks I am lazy.	.77
	20	She thinks I act ugly.	.73
	10	She gives us too much work.	.71
	6	She always picks on people.	.61
Factor 3: Fosterance of Self-Esteem	15	She likes me.	.79
	12	She thinks I am smart.	.72
	14	She thinks I can do a lot on my own.	.71
	11	She thinks I work hard.	.63
	19	She likes for me to help her.	.59

items", Item No. 3 and 7 will, in fact, separate out from Factor 1 (as they do at later ages along with Items 9 and 17) to form a separate factor. This differentiation which occurs progressively with age is roughly equivalent to the "love" and "competence" differentiation. The problem is that at the early ages "competence items", Item No. 3 and 7, which will form a fourth factor, also take a pure rapport item (16) with them into the factor: the separation is inexact and not as clean as that which occurs at grades 4-6 (See Tables 11 and 12 which included grades 1-6, and Table 13 for a comparison of the structures.) It cannot be stressed too much, however, that repeatedly in these data--regardless of age--the finding occurs that children's evaluation of a teacher (as here with "competence") is highly related to their belief about the teacher's attitude toward them (as here with "rapport"). The blue card "projective" items (Items 11-22) which were planned to give a separate estimate of the child's belief about the teacher's view of him, contained several items which were consistently and highly related to the child's more "objective" evaluation of the teacher herself (yellow card Items 1-10). Children's estimate of a teacher's competence appears to be highly associated with the children's belief about the teacher's view of <u>their</u> competence. Also, whether children think the teacher conducts a stimulating, pleasant

TABLE 11

Factor Analysis of SET II
Grades 1-6

72.4 Percent of the Variance Extracted by Four Roots as follows:

Factor 1	Factor 2	Factor 3	Factor 4
17.1	17.6	16.0	21.7

	Factor Loadings After Rotation to Simple Structure			
Variable (Item)	Factor 1	Factor 2	Factor 3	Factor 4
3	.19	-.38	.76	.17
4	.32	-.27	.25	.78
5	.22	-.19	.50	.70
6	-.09	.59	-.57	-.23
7	.26	-.15	.77	.27
8	.00	.84	-.05	-.20
9	.32	-.26	.44	.49
10	-.23	.72	-.27	.04
11	.61	-.02	-.18	.52
12	.73	-.17	.34	.41
13	.20	.20	.19	.81
14	.69	.10	.32	.21
15	.80	-.21	.08	.24
16	.22	-.18	.45	.70
17	.14	.11	.65	.32
18	.16	-.46	.26	.64
19	.74	-.31	.21	.07
20	-.02	.73	-.13	.35
21	.52	-.09	.27	.68
22	-.22	.76	.04	-.14

TABLE 12

Primary (Highest) Loadings
of SET II Items--Four Factor Structure
Grades 1-6

Factor:	Item No.	Statement	Loading
Rapport with Children (Factor 4, Table 9)	13	She thinks we are a smart class.	.81
	4	The kids like her.	.78
	5	She makes school fun.	.70
	16	She likes us kids.	.70
	21	She thinks I have good ideas.	.68
	18	She thinks that kids are good.	.64
	9	She listens to what we want.	.49
Interactional Competence (Factor 3, Table 9)	7	She helps us a lot.	.77
	3	We can tell how she wants things done.	.76
	17	She likes to teach.	.65
Unreasonable Negativity (Factor 2, Table 9)	8	She gets mad a lot.	.84
	22	She thinks I am lazy.	.76
	20	She thinks I act ugly.	.73
	10	She gives us too much work.	.72
	6	She always "picks on" people.	.59
Fosterance of Self-Esteem (Factor 1, Table 9)	15	She likes me.	.80
	19	She likes for me to help her.	.74
	12	She thinks I am smart.	.73
	14	She thinks I can do a lot on my own.	.69
	11	She thinks I work hard.	.61

class appears closely tied to what the children believe their teacher's view is of their intelligence; and children's report of their own self-esteem appears to be highly related to whether they believe the teacher likes them or not. The apparent conclusion of these data--and probably the most important conclusion--is that children seem to view teachers' behavior as a reaction of teachers to qualities in the children themselves. This line of thinking may indicate that remnants of early childhood ego-centric omnipotence remain in grades 1-6; but this attitude is also the beginning of wisdom. In other words, children at early ages appear to understand that they have a part in creating others' responses to themselves or, at least, they sense that some causal relationship exists between themselves and teachers' responses. The danger is probably that children can go too far in this assumption. This may account for children's apparent confusion when teachers are excessively and unreasonably negative (Factor 2, to be discussed) as children cannot "see" the basis for such negativity. Much damage to a child's self-concept can result from assuming too much responsibility for teachers' reactions.

In contrast to Factor 1, the cohesiveness of Factor 2 and Factor 3 is quite stable across all age groups tested. These two factors are concerned with the unreasonably negative

emotionality of the teacher and the degree to which the teacher fosters development of individual self-esteem. Evidently these factors remain virtually unchanged across grades 1-6 (Table 13) and are differentiated early.

A more detailed discussion of the content of the separate factors will follow in the next section.

TABLE 13

Comparison of Primary Factor Structures
Derived From Teacher Mean Scores

Item #	Item Statement	Four Factors, All 36 Teachers Grades 1-6: Primary Structure with 20 Items		Three Factors 28 Teachers Grades 1-3: Primary Structure with 20 Items		Four Factors, All 36 Teachers Grades 1-6: Primary Structure with 22 Items	
		Item:	Loading: Factor 4	Item:	Loading: Factor 1	Item:	Loading: Factor 4
Stimulating, Interactive Style:							
13	She thinks we are a smart class.	14	.81	14	.71	14	.79
4	The kids like her.	4	.78	4	.85	4	.76
16	She likes us kids.	17	.70	17	.80	17	.65
5	She makes school fun.	5	.70	5	.93	5	.62
21	She thinks I have good ideas.	22	.68	22	.67	22	.73
18	She thinks that kids are good.	19	.64	19	.65	19	.56
9	She listens to what we want.	9	.49				
		(12 loads here .52)		(12 loads here .20)		12	.72
Interactional Competence:					Factor 1		
7	She helps us a lot.	7	.77	7	.70	7	.79
3	We can tell how she wants things done.	3	.76	3	.65	8	.74
17	She likes to teach.	18	.65	18	.60	18	.67
9	She listens to what we want.	(9 loads here .44)		9	.66	9	.51

TABLE 13 (cont.)

		Four Factors, All 36 Teachers Grades 1-6: Primary Structure with 20 Items		Three Factors 28 Teachers Grades 1-3: Primary Structure with 20 Items		Four Factors, All 36 Teachers Grades 1-6: Primary Structure with 22 Items	
Item #	Item Statement	Item:	Loading:	Item:	Loading:	Item:	Loading:
Unreasonable Negativity:				Factor 2			
8	She gets mad a lot.	8	.84	8	.81	8	.84
22	She thinks I am lazy.	23	.76	23	.77	23	.77
20	She thinks I act ugly.	21	.73	21	.73	21	.75
10	She gives us too much work.	10	.72	10	.71	10	.70
6	She always picks on people.	6	.59	6	.61	6	.63
Fosterance of Self-Esteem:				Factor 3			
19	She likes for me to help her.	20	.74	20	.60	20	.63
12	She thinks I am smart.	13	.73	13	.72	13	.63
14	She thinks I can do a lot on my own.	15	.69	15	.71	15	.51
15	She likes me.	16	.64	16	.80	16	.73
11	She thinks I work hard.	12	.61	12	.63		
1	She teaches us a lot.	(omitted from analysis)		(omitted from analysis)		1	.78
2	She is a smart teacher.	(omitted from analysis)		(omitted from analysis)		2	.70

Factor 1: Stimulating, Interactive Style

This factor is determined by a group of items which suggest that a lively, mutually respectful sense of rapport exists between teacher and student. The teacher makes school fun, the teacher and children like each other, and the teacher has a high regard for the children's intelligence.[4] The flavor is essentially that of an egalitarian "with-it-ness" such as can be observed often in the behavior of adults who direct children's T.V. programs--for example, Sesame Street. The children and the teacher are mutually engaged in a pleasant, stimulating relationship in which the teacher does not "talk down" to the children but expects from them an intelligent degree of understanding. This is the sense of "love" or rapport which is suggested by these items. This sense of rapport is definitely compatible with that discussed extensively in the Review of the Literature as a "generous, non-egocentric view of children" by the teacher.

The second "set" of items which loads upon Factor 1 at grades 1-3 are the items which chiefly concern the teacher's

[4]The relationship of teacher rapport and the teacher's supposed view of the children's intelligence is so strong that even Item #19, which is a self-esteem item stated in the first person, "crosses over" to load more heavily on Factor 1 than it does on Factor 3. The item is classified on the final form of the SET II upon Factor 3, where it logically belongs (and will appear with increased age) and where it also loads highly.

competence, specifically the success with which the teacher interacts and communicates with the class. These items are highly related to the rapport items on Factor 1 (discussed above) and, in fact, maintain much of this relationship even with increasing age and their eventual separation into another discrete factor. These items suggest that the teacher can both communicate her own intentions and listen empathetically to the students; she helps the students a lot; and she likes to teach. These items definitely concern the act of teaching in its most attractive form: the teacher and students are communicating, the teacher is offering appropriate aid where the "gaps" exist, and the teacher gives the appearance of deriving personal satisfaction from her part in the process. These items measure the same "with-it" quality that the rapport items measure, but they are less emotional and evaluative and more competency-based.

All items on Factor 1 suggest a high degree of successful student-teacher interaction. The "rapport" and "competence" items will separate with increasing age, and they may even be "viewed" separately on protocols of teachers in grades 1-3 for any clinical insights they may offer. But there appears to be only limited statistical justification for expecting these dimensions to exist separately in the minds of most first-third

graders. That is the reason that Factor 1 is entitled simply "Stimulating Interactive Style" on the SET II when it is used with children below the third grade level. To view the quality being measures as "simple", however, would be quite misleading.

Factor II: Unreasonable Negativity

The hypothesis was entertained in the Review of the Literature section (this manual) that teacher rating instruments do not tap sufficiently the child's strong reaction to excessive emotional negativity in a teacher's behavior. Results of the SET II data analysis suggest that this factor is even stronger than anticipated. (The item loadings on this factor are presented in Tables 9 and 10.)

"She gets mad a lot" determines this factor. Also primary on Factor 2 are two items which express the child's belief that the teacher disapproves of his behavior, followed by two items which indicate that the teacher behaves in a punitive manner with the class. Not to be ignored is the -.51 loading on this factor of "We can tell how she wants things done". Webster (Merrian-Webster Seventh New Collegiate Dictionary, 1972) discusses the term unreasonable as connoting "guidance by some force other than reason (as self will, rage) that makes one deficient in good sense" (see irrational, above reference).

This seems precisely the quality children are reacting to in Factor 2. The teacher is guided by some force other than reason, a force which is not apparent to the children. The teacher is inscrutable to them. She explodes and punishes, but they cannot ascertain what it is that she wants. They indicate that she believes them to be lazy and to act ugly. But behaving in a positive, competent manner does not seem to be related to the teacher's negativity (the lowest loadings on this factor are for the three items of "She thinks I have good ideas," "She thinks I work hard," and "She thinks I can do a lot on my own").

The teacher who is receiving high scores on this factor is clearly a mystery to children. Probably the reason that children react so strongly to this type of negative emotionality is that they cannot comprehend any rational basis for its occurence. For this reason, the factor has been named "Unreasonable Negativity". It is not to be expected, however, that all openly emotionally-reactive teachers will necessarily score very high on this factor (though that remains to be discovered). Children are apparently tapping a more fearsome quality than simple emotionality on Factor 2: they see the anger, they feel the disapproval, and they incur the punishment; but they don't know what the impetus for such teacher behavior can be.

Factor III: Fosterance of Self-Esteem

With the exception of the first item, all items which load heavily on Factor 3 are items that clearly relate to personal competence and self-esteem: She thinks "I am smart..." ; "I can do a lot on my own...."; "I work hard"; "She likes for me to help her "; and "She thinks I have good ideas ". The primary item on this factor, however, is "She likes me ". The implied causal link is probably, "She likes me because "I am smart and competent", etc. (What is more, this relation strengthens rather than weakens with age.) As in Factor 1, a definite relationship exists factorially between teacher behavior as the child observes it and what the child believes are his teacher's evaluations of his own intelligence and his competence. It is for this reason, plus the fact that this factor analysis is based upon teacher mean scores rather than aggregate individual student responses, that student self-esteem (as reported on the "blue card" items) is viewed as an interactive function of teacher behavior and the student's characteristics. In other words, differences in teacher mean scores on what might be viewed as purely personal, "projective" items cannot be accounted for by the nature of the students themselves. Something is going on in a certain teacher's class when her students as a group

feel better or worse about themselves than other students feel about themselves in groups. This "something" is proposed as being the degree to which the teacher fosters the development of an individual sense of self-esteem in children. The first item on this factor is viewed as quite supportive of this view.

Final Form of the Experimental SET II

The final form of the SET II abandons any attempt to separate the student's more "objective" evaluation of the teacher from the students more "projective" attempt to estimate his teacher's supposed attitudes toward himself. That there appears to be no such separation in a child's evaluations (as has been discussed) is an important finding. Instead, the final form of the SET II is presented in the three-factor format which has been discussed. All three factors involve interaction between teacher and student to some degree.

During the analyses, two items were dropped from the original twenty-two items, though Item No. 1 is retained for its administrative function as a "warm-up" item. Both items provided little information as the standard deviation of responses to these items was markedly lower than for the responses to any other items (See Table 14 & 15). Also, and probably because of this,

TABLE 14

Means and Standard Deviations:
Pooled Student Responses All Items,
Grades 1-3 and Grades 4-6

	Grades 1-3 only:		Grades 4-6 Inc.:		
Item:	M	SD*	M	SD*	Item Statement:
1	1.04	.198	1.11	.313	She teaches us a lot.
2	1.04	.186	1.08	.269	She is a smart teacher.
3	1.27	.445	1.25	.430	We can tell how she wants things done.
4	1.10	.297	1.29	.453	The kids like her.
5	1.17	.379	1.27	.443	She makes school fun.
6	1.79	.404	1.82	.382	She always picks on people.
7	1.09	.282	1.11	.315	She helps us a lot.
8	1.71	.455	1.57	.495	She gets mad a lot.
9	1.37	.489	1.27	.443	She listens to what we want.
10	1.67	.470	1.55	.497	She gives us too much work.
11	1.17	.375	1.37	.484	She thinks I work hard.
12	1.11	.316	1 31	.461	She thinks I am smart.
13	1.09	.290	1.19	.394	She thinks we are a smart class.
14	1.15	.361	1.25	.436	She thinks I can do a lot on my own.
15	1.07	.258	1.21	.406	She likes me.
16	1.10	.303	1.13	.336	She likes us kids.
17	1.07	.249	1.05	.208	She likes to teach.
18	1.12	.329	1.20	.402	She thinks that kids are good.
19	1.15	.361	1.24	.429	She likes for me to help her.
20	1.82	.388	1.74	.437	She thinks I act ugly.
21	1.21	.405	1.38	.486	She thinks I have good ideas.
22	1.82	.388	1.77	.422	She thinks I am lazy.

*Standard Deviation

T= Score of 1

F= Score of 2

TABLE 15

Standardized Means: The Teacher Item Means

All 36 Teachers, All 22 Items

Item #:	Item Statement:	M	S.D.*
1	She teaches us a lot.	1.15	.144
2	She is a smart teacher.	1.08	.096
3	We can tell how she wants things done.	1.23	.136
4	The kids like her.	1.16	.181
5	She makes school fun.	1.19	.198
6	She always "picks on" people.	1.81	.150
7	She helps us a lot.	1.11	.102
8	She gets mad a lot.	1.68	.186
9	She listens to what we want.	1.27	.201
10	She gives us too much work.	1.62	.163
11	She thinks I work hard.	1.19	.120
12	She thinks I am smart.	1.18	.147
13	She thinks we are a smart class.	1.14	.136
14	She thinks I can do a lot on my own.	1.18	.115
15	She likes me.	1.14	.126
16	She likes us kids.	1.13	.123
17	She likes to teach.	1.07	.079
18	She thinks that kids are good.	1.19	.164
19	She likes for me to help her.	1.17	.122
20	She thinks I act ugly.	1.76	.146
21	She thinks I have good ideas.	1.23	.151
22	She thinks I am lazy.	1.78	.148

Note: T = Score of 1
F = Score of 2

*Standard Deviation

both items tended to distort the logical structure of many factor analyses performed with the two items included in the analysis. Furthermore, a number of teachers found Item No. 2 ("She is a smart teacher") to be offensive; and there was some indication that students at times used the item punitively rather than logically. Finally, it makes sense that children who have spent so few years in school demonstrate little variation in their responses to these two items. Their lack of experience with various teachers makes it difficult for them to tell whether a certain teacher is smart and/or teaches "a lot" or not.

At older ages, responses to these two items become more diverse. But even then, the items do not load so highly with the teacher-competence items on Factor 1 as they do with the student's own self-esteem and competence on Factor 3. Taken together, the reasons above seemed sufficient for discarding the items as sources of useful information about teachers.

The final form of the Experimental SET II is presented in Appendix D.

CHAPTER V

USING THE SET II

The Experimental Version of the SET II is now available for use in research or teacher training programs. In the latter case, The SET II is viewed as appropriate for confrontation (or feedback) counseling purposes as well as for evaluation. Professional persons may arrange to obtain copies of the SET II for use in Grades 1-3 by using the following address:

SET II Order
Dissemination Department
Research & Development Center
for Teacher Education
Educational Annex
University of Texas
Austin, Texas 78703

Persons desiring to use the SET II with a population in Grades 4-6 may use the one-page form found in Appendix L. All persons who use the Experimental SET II are requested to share their data with The Research and Development Center for Teacher Education. This request is made in order that norms may be developed for teaching behavior around the country. At present, such norms are not available.

In the absence of norms based upon a widely-distributed teacher population, groups who wish to use the SET II will presently need to develop their own norms. The normative data

contained in this manual (Table 16) may also be used as a reference base when it is impossible for a school group to develop its own norms (for example, in very small schools with few teachers). In such cases, extreme caution should be used in drawing conclusions, however, since the data contained in this manual are based upon responses to 36 teachers at most. The primary purpose of the Pilot SET II study was the investigation of students' reliability of response. As more data are gathered by further testing of the SET II, both normative findings and validation of the factor structure of the instrument will be pursued.

Using the SET II for Feedback Purposes

The most common use of the SET II will probably be that of providing feedback to teachers so that they can evaluate their own effectiveness. The typical way to do this will be to discuss with the teacher her own profile of scores. Information for the profile will need to be derived in the following manner:

A Teacher Tally Sheet is presented in Appendix E. Tally all the responses received by one teacher on a single tally sheet.[5] The total number of positive and negative responses to each item will be recorded. The total negative scores and positive scores for all items in a single scale (based on the factor

[5] These responses come from the individual students' tally sheets. (See Appendix I.)

TABLE 16

Standardized Scales
20 Item SET II, 36 Teachers

Scale 1: Stimulating Interactive Style (R+I)

10.4764 = Mean
1.0789 = Standard Deviation
.1798 = Standard Error of Mean
3.6020 = Skewness Z, (P = .0006)

Raw Score	Frequency	Percentage	Percentile	Standard	Normal
9.19	3	8	4	38	33
9.30	2	6	11	39	38
9.51	2	6	17	41	40
9.62	1	3	21	42	42
9.73	2	6	25	43	43
9.83	2	6	31	44	45
10.05	2	6	36	46	46
10.15	1	3	40	47	48
10.26	5	14	49	48	50
10.37	1	3	57	49	52
10.47	1	3	60	50	52
10.58	2	6	64	51	54
10.68	1	3	68	52	55
10.79	1	3	71	53	55
10.90	1	3	74	54	56
11.00	2	6	78	55	58
11.11	1	3	82	56	59
11.22	1	3	85	57	60
11.32	1	3	88	58	62
11.96	2	6	92	64	64
13.24	1	3	96	76	67
14.09	1	3	99	84	72

TABLE 16 (cont.)

Subscale: Rapport (or Scale 4, above third grade)

5.7986 = Mean
.7172 = Standard Deviation
.1195 = Standard Error of Mean
3.9207 = Skewness Z, (P = .0003)

Raw Score	Frequency	Percentage	Percentile	Standard	Normal
5.02	2	6	3	39	31
5.09	1	3	7	40	35
5.16	5	14	15	41	40
5.23	2	6	25	42	43
5.37	2	6	31	44	45
5.44	1	3	35	45	46
5.51	1	3	37	46	47
5.58	4	11	44	47	49
5.65	3	8	54	48	51
5.72	3	8	63	49	53
5.79	1	3	68	50	55
6.08	1	3	71	54	55
6.15	2	6	75	55	57
6.22	3	8	82	56	59
6.36	1	3	88	58	62
6.85	1	3	90	65	63
7.07	1	3	93	68	65
7.77	1	3	96	78	67
8.13	1	3	99	83	72

TABLE 16 (cont.)

Subscale: Interactional Competence (or Scale 1, above third grade)

4.6778 = Mean
.4284 = Standard Deviation
.0714 = Standard Error of Mean
3.7100 = Skewness Z, (P = .0005)

Raw Score	Frequency	Percentage	Percentile	Standard	Normal
4.04	2	6	3	35	31
4.12	1	3	7	37	35
4.21	3	8	13	39	38
4.29	1	3	18	41	41
4.42	2	6	22	44	42
4.46	2	6	28	45	44
4.50	1	3	32	46	45
4.55	2	6	36	47	46
4.59	1	3	40	48	48
4.63	2	6	44	49	49
4.67	6	17	56	50	51
4 .72	2	6	67	51	54
4.80	2	6	72	53	56
4.88	1	3	76	55	57
4.93	2	6	81	56	59
5.01	1	3	85	58	60
5.10	1	3	88	60	62
5.14	2	6	92	61	64
5.48	1	3	96	69	67
6.32	1	3	99	89	72

TABLE 16 (cont.)

Scale 2: Unreasonable Negativity

8.6608 = Mean
.6385 = Standard Deviation
.1064 = Standard Error of Mean
-.5599 = Skewness Z, (P = .5827)

Raw Score	Frequency	Percentage	Percentile	Standard	Normal
7.27	1	3	1	28	28
7.65	1	3	4	34	33
7.77	2	6	8	36	36
7.90	1	3	13	38	38
7.96	1	3	15	39	40
8.03	1	3	18	40	41
8.09	1	3	21	41	42
8.15	1	3	24	42	43
8.22	2	6	28	43	44
8.28	3	8	35	44	46
8.34	1	3	40	45	48
8.53	1	3	43	48	48
8.66	1	3	46	50	49
8.72	1	3	49	51	50
8.78	2	6	53	52	51
8.84	1	3	57	53	52
8.91	1	3	60	54	52
9.03	2	6	64	56	54
9.10	3	8	71	57	55
9.16	1	3	76	58	57
9.22	1	3	79	59	58
9.35	1	3	82	61	59
9.41	1	3	85	62	60
9.47	2	6	89	63	62
9.54	2	6	94	64	66
9.60	1	3	99	65	72

TABLE 16 (cont.)

Scale 3: Fosterance of Self Esteem

7.0944 = Mean
.6343 = Standard Deviation
.1057 = Standard Error of Mean
2.1760 = Skewness Z, (P =.0277)

Raw Score	Frequency	Percentage	Percentile	Standard	Normal
6.21	2	6	3	36	31
6.28	1	3	7	37	35
6.40	3	8	13	39	38
6.46	2	6	19	40	41
6.65	3	8	26	43	44
6.71	1	3	32	44	45
6.84	4	11	39	46	47
6.90	2	6	47	47	49
6.96	1	3	51	48	50
7.03	1	3	54	49	51
7.15	2	6	58	51	52
7.28	1	3	63	53	53
7.34	2	6	67	54	54
7.40	3	8	74	55	56
7.46	1	3	79	56	58
7.53	1	3	82	57	59
7.71	1	3	85	60	60
7.78	1	3	88	61	62
8.03	1	3	90	65	63
8.09	1	3	93	66	65
8.59	1	3	96	74	67
8.84	1	3	99	78	72

structures described earlier) may then be added together. What results is one total score for each scale. (Use the four-factor form for Grades 4-6.)

Each of a teacher's scale totals (there will be 3 or 4) can then be divided by the number of students in the teacher's class who responded to the instrument. This operation yields the teacher's mean (or average) scale scores for each scale.

Normative data for an entire teacher group[6] should then be generated by using all the individual teachers' mean scale scores as the scores to be manipulated. From these teacher mean scores, a teacher group mean and a teacher group standard deviation (i.e., "norms") can be derived for each factor. (See Appendix K[7])

These teacher group norms will be posted in parentheses on the Teacher Profile Form (See Teacher Profile Form, Appendix G). After the teacher's own scale scores have also been posted on the profile, in the open blanks, the completed profile can used to provide feedback to the teacher herself.

---

[6] "Teacher group" = The total number of teachers being tested, which should be at least 25-30.

[7] If professional help is not available locally to complete the generation of school group norms, see Appendix K for a method of computing these normative scores to be placed in the parentheses on the Profile Forms.

Use of the SET II in Research

Scoring the SET II is essentially the same as described above. In research activities, of course, computer programming is usually available to perform the tasks. The use of teacher mean scores in data analyses is strongly suggested so that the emergence of any logical structure which may exist in students' patterns of responses is best encouraged.

At present, Item No. 2 is not used in the card system. The former Item 1 ("She teaches us a lot") is used for administrative purposes but not scored. As a next step in development of the SET II, two new items are now proposed to replace old Items No. 1 and 2. Both of these items are to be administered, to be recorded as part of the teacher's tally sheet, but not to be included in scoring. It is hoped both items will eventually bolster the "interactional competence" dimension; and when sufficient data have been gathered to determine whether the items are functioning as intended, they will be incorporated into the scoring of the SET II or replaced, whichever is indicated. The new items are:

No. 2: She makes what we learn interesting.

No. 23: She is nice when we make mistakes.

Communication regarding the SET II is invited. Please contact Dr. Ruth A. Haak at the Research and Development Center for Teacher Education, The University of Texas, Austin, Texas 78712.

APPENDIX A

She thinks that kids are good.

She likes to teach.

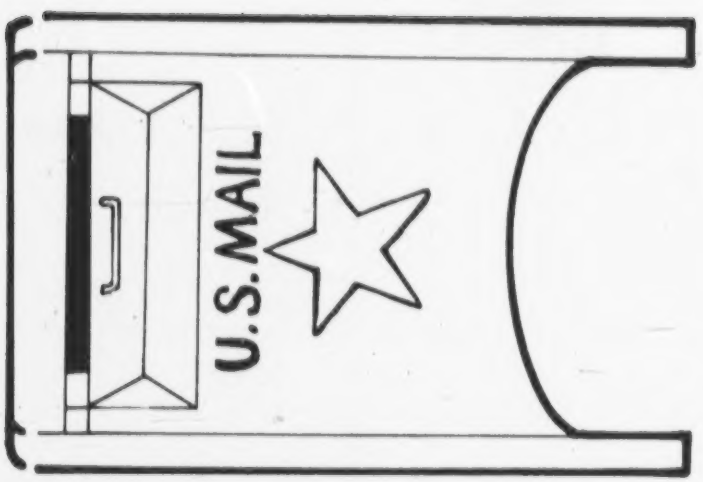
She thinks I have good ideas.

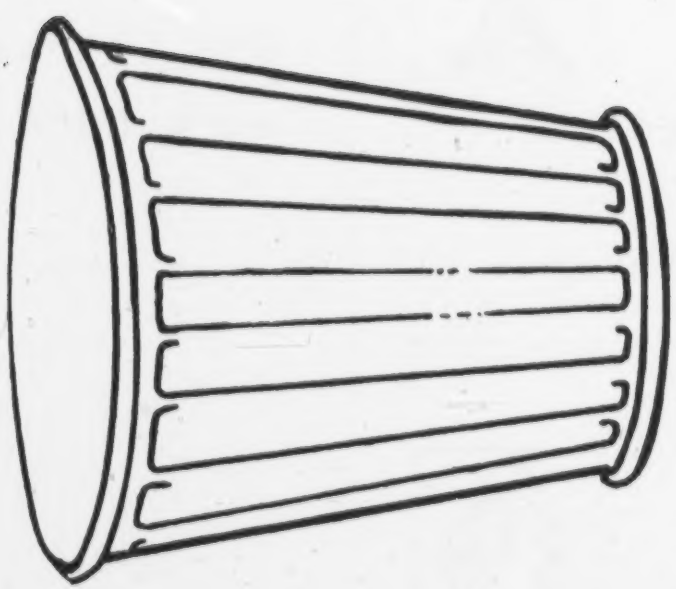
She thinks I work hard.

She likes me.

She likes us kids.

APPENDIX B





APPENDIX C

Besides the factor analyses which are presented in this manual, the following analyses were performed and are available from the author on request:

1. Factor Analysis: Grades 2 & 3 (All)
2. Factor Analysis: Grades 1-3 White
3. Factor Analysis: Grades 1-3, Ability 2-3, Mexican-American
4. Factor Analysis: Grades 1-3, Ability 2-3, Black
5. Factor Analysis: First 10 Items, Grades 2-3, Black
6. Factor Analysis: First 10 Items, Grades 2-3, Mexican-American
7. Factor Analysis: First 10 Items, Grades 2-3, White
8. Factor Analysis: First 10 Items, Grades 4-6, Mexican-American
9. Factor Analysis: Last 12 Items, Grades 2-3, Black
10. Factor Analysis: Last 12 Items, Grades 2-3, Mexican-American
11. Factor Analysis: Last 12 Items, Grades 2-3, White
12. Factor Analysis: Last 12 Items, Grades 4-6, Mexican-American

APPENDIX D

The Factor Structure of the SET II

I. Stimulating, Interactive Style:

Item	Subscale: Rapport
5	She makes school fun.
4	The kids like her.
16	She likes us kids.
13	She thinks we are a smart class.
18	She thinks kids are good.

	Subscale: Interactional Competence
7	She helps us a lot.
9	She listens to what we want.
3	We can tell how she wants things done.
17	She likes to teach.

II. Unreasonable Negativity:

Item	
8	She gets mad a lot.
22	She thinks I am lazy.
20	She thinks I act ugly.
10	She gives us too much work.
6	She always "picks on" people.

III. Fosterance of Self-Esteem:

Item	
15	She likes me.
12	She thinks I am smart.
14	She thinks I can do a lot on my own.
11	She thinks I work hard.
19	She likes for me to help her.
21	She thinks I have good ideas.

*For clinical or counseling use only at Grades 1-3.

APPENDIX E

Teacher Tally Sheet

Instructions: All the individual student's scoring records for one teacher should be posted on this sheet.

Teacher_____Grade_____Date_____

**N=_____

		True (T=1)		False (F=2)	
		Tally	Score	Tally	Score
I.	Stimulating, Interaction Style:				
	Item				
5	She makes school fun.				
4	The kids like her.				
15	She likes us kids.				
13	She thinks we are a smart class.				
18	She thinks kids are good.				

*Subscore: Total Rapport (T)= (F)=

*Subtotal:_____

÷ N = _____

(Scale Mean Score)

7	She helps us a lot.				
9	She listens to what we want.				
3	We can tell how she wants things done.				
17	She likes to teach.				

*Subscore: Total Interactional Competence (T)= (F)=

*Subtotal:_____

÷ N = _____

(Scale Mean Score)

Scale Total_____

÷ N =_____

(Scale Mean Score)

T F

New Items:

2 She makes what we learn interesting. (tally only) _____

23 She is nice when we make mistakes. (tally only)_____

APPENDIX E (cont.)

True (T=1) False (F=2)

II. Unreasonable Negativity

Item:

8 She gets mad a lot.
22 She thinks I am lazy.
20 She thinks I act ugly.
10 She gives us too much work.
6 She always "picks" on people.

Tally	Score	Tally	Score
T=		F=	

Scale Total _____

÷N _____

(Scale Mean Score)

III. Fosterance of Self-Esteem

Item:

15 She likes me.
12 She thinks I am smart.
14 She thinks I can do a lot on my own.
11 She thinks I work hard.
19 She likes for me to help her.
21 She thinks I have good ideas.

T=		F=	

Scale Total _____

÷N _____

(Scale Mean Score)

**N= number of students who rated this teacher.

* For clinical or counseling use only at Grades 1-3. At Grades 4 and above, use as separate scale scores. (Rapport becomes the name of Scale 4, Interactional Competence the name of Scale 1.)

EXAMPLE

APPENDIX F

Teacher Tally Sheet

Instructions: All the individual student's scoring records for one teacher should be posted on this sheet.

Teacher Doe, Molly Grade 2 Date 12-30-72

**N= 20

		True (T=1)		False (F=2)	
		Tally	Score	Tally	Score
I.	Stimulating, Interaction Style:				
	Item				
5	She makes school fun.	𝍸 𝍸 𝍸	15	𝍸	10
4	The kids like her.	𝍸 𝍸 𝍸	15	𝍸	10
15	She likes us kids.	𝍸 𝍸 𝍸 II	17	III	6
13	She thinks we are a smart class.	𝍸 𝍸 𝍸 I	16	IIII	8
18	She thinks kids are good.	𝍸 𝍸 III	13	𝍸 II	14
	*Subscore: Total Rapport		(T) = 76		(F) = 48
				*Subtotal: 124	
				÷ N =	
				(Scale Mean Score)	
7	She helps us a lot.	𝍸 𝍸 IIII	14	𝍸 I	12
9	She listens to what we want.	𝍸 𝍸 𝍸	15	𝍸	10
3	We can tell how she wants things done.	𝍸 𝍸 𝍸 III	18	II	4
17	She likes to teach.	𝍸 III	8	𝍸 𝍸 II	24
	*Subscore: Total Interactional Competence		(T) = 55		(F) = 50
				*Subtotal: 105	
				÷ N =	
				(Scale Mean Score)	
				Scale Total 229	
				÷ N = 11.45	
				(Scale Mean Score)	

New Items:

		T	F
2	She makes what we learn interesting. (tally only)	𝍸 𝍸 𝍸	𝍸
23	She is nice when we make mistakes. (tally only)	𝍸 𝍸 II	𝍸 III

APPENDIX F (cont.)

II. Unreasonable Negativity:

Item		True (T=1) Tally	True (T=1) Score	False (F=2) Tally	False (F=2) Score
8	She gets mad a lot.	IIII	4	||| ||| ||| I	32
22	She thinks I am lazy.	III	3	||| ||| ||| II	34
20	She thinks I act ugly.	II	2	||| ||| ||| III	36
10	She gives us too much work.	IIII	4	||| ||| ||| I	32
6	She always "picks on" people.	I	1	||| ||| ||| IIII	38
		(T) =	14	(F) =	172
	Scale Total				186
	÷ N =				9.30
	(Scale Mean Score)				

III. Fosterance of Self-Esteem:

Item					
15	She likes me.	||| ||| ||| ||| |||	20		0
12	She thinks I am smart.	||| ||| ||| |||	16	IIII	8
14	She thinks I can do a lot on my own.	||| ||| ||| ||| |||	20		0
11	She thinks I work hard.	||| ||| ||| ||| II	17	III	6
19	She likes for me to help her.	||| ||| ||| ||| |||	19	I	2
21	She thinks I have good ideas.	||| ||| ||| ||| II	17	III	6
		(T) =	109	(F) =	22
	Scale Total				131
	÷ N =				6.55
	(Scale Mean Score)				

**N - Number of students who rated this teacher.

*For clinical or counseling use only at Grades 1-3. At Grades 4 and above, use as separate scale scores. (Rapport becomes Scale 4, Interactional Competence remains Scale 1.)

APPENDIX G

Name_____Grade_____Date_____

Teacher's three scores may be placed upon this profile to give a picture of her comparative strengths and weaknesses.

Teacher Profile (For Grades 1-3)

	Stimulating, Interactive Style	Unreasonable Negativity	Fosterance of Self-Esteem
+3S	( )*	( )	( )
+2S	( )	( )	( )
+1S	( )	( )	( )
M	( )	( )	( )
-1S	( )	( )	( )
-2S	( )	( )	( )
-3S	( )	( )	( )
Group Mean	_____	_____	_____
Standard Deviation	_____	_____	_____

Post norm figures from high to low ↑

*Figures are to be filled in from local norms, based upon a sufficient number of subjects.

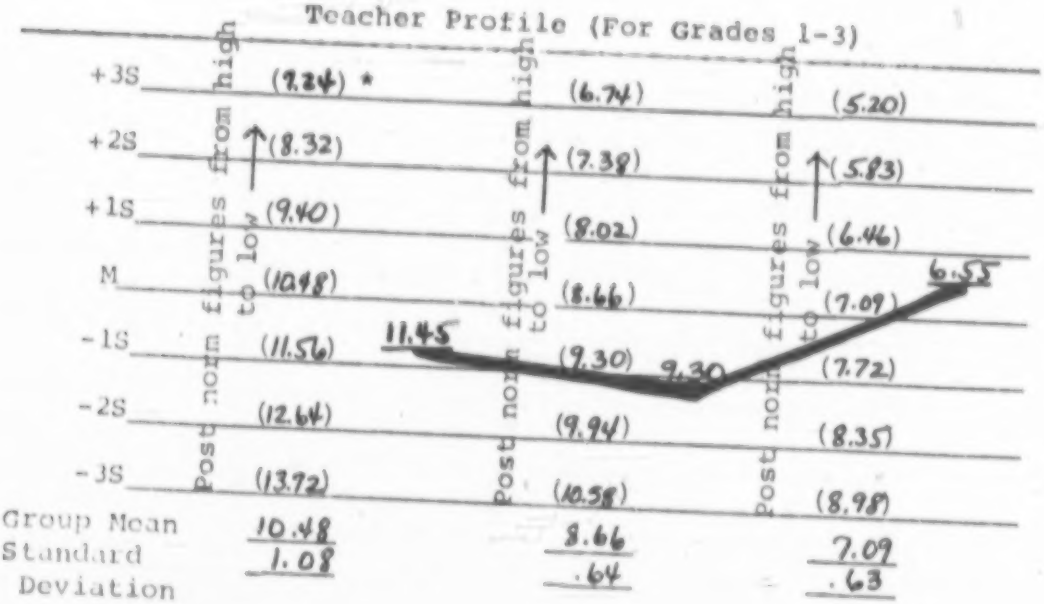
NOTE: Since 1=T and 2=F in scoring the SET II, the meaning of each scale score must be interpreted carefully. On Scales 1 and 3, a high mathematical score has a negative connotation. On Scale 2, a high mathematical score has a positive connotation. To avoid unnecessary confusion in using this SET II Profile for counseling purposes, post the norm figures for ALL SCALES from high at the bottom of the profile to low on top of the profile. This makes the interpretation of all scales easier, as each scale score which appears above the mean then carries with it a connotation that this is a "high" score on that scale--which is the usual case in test interpretation.

EXAMPLE

APPENDIX H

Name Doe, Molly Grade 2 Date 12-30-72

Teacher's three scores may be placed upon this profile to give a picture of her comparative strengths and weaknesses.



*Figures are to be filled in from local norms, based upon a sufficient number of subjects.

NOTE: Since 1=T and 2=F in scoring the SET II, the meaning of each scale score must be interpreted carefully. On Scales 1 and 3, a high mathematical score has a negative connotation. On Scale 2, a high mathematical score has a positive connotation. To avoid unnecessary confusion in using this SET II Profile for counseling purposes, post the norm figures for ALL SCALES from high at the bottom of the profile to low on top of the profile. This makes the interpretation of all scales easier, as each scale score which appears above the mean then carries with it a connotation that this is a "high" score on that scale--which is the usual case in test interpretation.

APPENDIX I

SET II Pupil's Scoring Record

Instructions:

Record the way the student sorted his cards upon this sheet as the cards are pulled from the sorting envelope.

Student's Name_____Grade____ Date_____

I. Stimulating, Interactive Style:

Item		True	False
5	She makes school fun.	(T=1)____	(F=2)____
4	The kids like her.	(T=1)____	(F=2)____
16	She likes us kids.	(T=1)____	(F=2)____
13	She thinks we are a smart class.	(T=1)____	(F=2)____
18	She thinks kids are good.	(T=1)____	(F=2)____
	*Subscore: Rapport	(T=1)____	(F=2)____
7	She helps us a lot.	(T=1)____	(F=2)____
9	She listens to what we want.	(T=1)____	(F=2)____
3	We can tell how she wants things done.	(T=1)____	(F=2)____
17	She likes to teach.	(T=1)____	(F=2)____
	*Subscore: Interactional Competence	(T=1)____	(F=2)____

Scale Total: (T) + (F) = _____

New Items:		T	F
2	She makes what we learn interesting.	(tally only)	____
23	She is nice when we make mistakes.	(tally only)	____

APPENDIX I (cont.)

II. Unreasonable Negativity:

Item		
9	She gets mad a lot.	(T=1)____ (F=2)____
22	She thinks I am lazy.	(T=1)____ (F=2)____
20	She thinks I act ugly.	(T=1)____ (F=2)____
10	She gives us too much work.	(T=1)____ (F=2)____
4	She always "picks on" people.	(T=1)____ (F=2)____

Scale Total: (T) + (F) = _____

III. Fosterance of Self-Esteem:

Item		
15	She likes me.	(T=1)____ (F=2)____
12	She thinks I am smart.	(T=1)____ (F=2)____
14	She thinks I can do a lot on my own.	(T=1)____ (F=2)____
11	She thinks I work hard.	(T=1)____ (F=2)____
19	She likes for me to hel, her.	(T=1)____ (F=2)____
21	She thinks I have good ideas.	(T=1)____ (F=2)____

Scale Total: (T) + (F) = _____

*For clinical or counseling use only at Grades 1-3. At Grades 4 and above, use as separate scale scores. (Rapport becomes Scale 4, Interactional Competence remains Scale 1.)

APPENDIX J

SET II Pupil's Scoring Record
(Example)

Instructions:

Record the way the student sorted his cards upon this sheet as the cards are pulled from the sorting envelope.

Student's Name Abner Student Grade 2 Date 12-30-72

I. Stimulating, Interactive Style:

Item		True	False
5	She makes school fun.	(T=1) ✓	(F=2)
4	The kids like her.	(T=1)	(F=2) ✓
16	She likes us kids.	(T=1)	(F=2) ✓
13	She thinks we are a smart class.	(T=1) ✓	(F=2)
18	She thinks kids are good.	(T=1) ✓	(F=2)
	*Subscore: Rapport	(T=1) 3	(F=2) 4
7	She helps us a lot.	(T=1) ✓	(F=2)
9	She listens to what we want.	(T=1)	(F=2) ✓
3	We can tell how she wants things done.	(T=1) ✓	(F=2)
17	She likes to teach.	(T=1)	(F=2) ✓
	*Subscore: Interactional Competence	(T=1) 2	(F=2) 4
	Scale Total:	(T) 5 + (F) 8	= 13

New Items:

		T	F
2	She makes what we learn interesting. (tally only)	✓	
23	She is nice when we make mistakes. (tally only)	✓	

2

APPENDIX J (cont.)

II. Unreasonable Negativity:

Item			
8	She gets mad a lot.	(T=1)____	(F=2)✓
22	She thinks I am lazy.	(T=1)✓	(F=2)____
20	She thinks I act ugly.	(T=1)____	(F=2)✓
10	She gives us too much work.	(T=1)✓	(F=2)____
6	She always "picks on" people.	(T=1)____	(F=2)✓

Scale Total: (T) 2 + (F) 6 = 8

III. Fosterance of Self-Esteem:

Item			
15	She likes me.	(T=1)✓	(F=2)____
12	She thinks I am smart.	(T=1)✓	(F=2)____
14	She thinks I can do a lot on my own.	(T=1)____	(F=2)✓
11	She thinks I work hard.	(T=1)____	(F=2)✓
19	She likes for me to help her.	(T=1)✓	(F=2)____
21	She thinks I have good ideas.	(T=1)✓	(F=2)____

Scale Total: (T) 4 + (F) 4 = 8

*For clinical or counseling use only at Grades 1-3. At Grades 4 and above, use as separate scale scores. (Rapport becomes Scale 4, Interactional Competence remains Scale 1.)

APPENDIX K

Computing Norms

In order to compare a teacher's performance to some set of standards, it is necessary to know what the standards are. In psychological testing, such standards are referred to as "norms". The two basic norms which are needed for use in interpreting performance on the SET II scales are a total group mean (the average score for the entire group) and a total group standard deviation (a figure which shows how much spread there is in the scores; this spread is measured at standard interval on a typical "bell curve" distribution and is called the "standard deviation").

With the SET II, a mean and standard deviation is needed for the performance of the group on each scale. This means that with data from third grade or below, three scale means and three scale standard deviations need to be computed for the entire group. If the data originates from fourth graders or above, four sets of such scores are needed. (See structure of the Experimental SET II, Table 13.)

If computational services are not available to a school, the following operations will provide a way to derive the necessary figures.

Factor Means

First, complete a Tally Sheet for each teacher (See Appendix F). Carry out all steps on the tally sheet. This will result in a set of scale scores for each teacher. Be sure all the teacher tally sheets are complete. To establish norms, now begin to work with one scale at a time. In other words, compute all figures for all teachers on Scale 1, then repeat for Scale 2, etc. The following example illustrates the process of establishing the group mean for Scale 1:

Example: Scale 1 (6 Teachers*)

Teacher:	Scale 1 Total Score:
Alice A.	17.49
Betty B.	15.40
Cora C.	14.32
Don D.	12.75
Elbert E.	11.63
Fred F.	9.41
Grand Total	81.00

After all teachers' Scale 1 scores have been added together, divide this grand total by the number of teachers involved:

$$81.00 \div 6 = 13.50$$

The group mean score for Factor 1 is 13.50.

Repeat this process for Scale 2 and Scale 3 (and Scale 4, if above third grade).

*Six is too few to compute reliable norms and is used here only for illustration.

Standard Deviations

In computing the standard deviations for each scale, again work with one scale at a time. The following steps, which are illustrated using fictitious data in the table below, can be followed in computing each standard deviation.

Step 1. List all teachers' scale scores for the scale in question.

Step 2. From each teacher's scale score subtract the group mean that you have just computed. This procedure yields a "deviation from the mean" score for each teacher. (Some of the deviations will be negative. All minus signs can be ignored.)

Step 3. Square all of the deviations obtained in Step 2. This gives a "squared deviation" score for each teacher.

Step 4. Sum all the squared deviations obtained in Step 3.

Step 5. Divide the sum obtained in Step 4 by the number of teachers. (E.g. If there are 34 teacher scale scores listed in Step 1, divide the sum obtained in Step 4 by 34.)

Step 6. Find the square root of the number obtained in Step 5. This is the standard deviation. (Consult a table of square roots to derive this figure.)

Illustration

Teachers	Scale Scores	Deviations	Squared Deviations
1. Alice A.	17.49	3.99	15.92
2. Betty B.	15.40	1.90	3.61
3. Cora C.	14.32	0.82	.67
4. Don D.	12.75	-0.75	.56
5. Elbert E.	11.63	-1.87	3.50
6. Fred F.	9.41	-4.09	16.73
Total	81.00		40.99

Group Mean = $81 \div 6 = 13.50$ Standard Deviation = $\sqrt{\frac{40.99}{6}} = \sqrt{6.83} = 2.61$

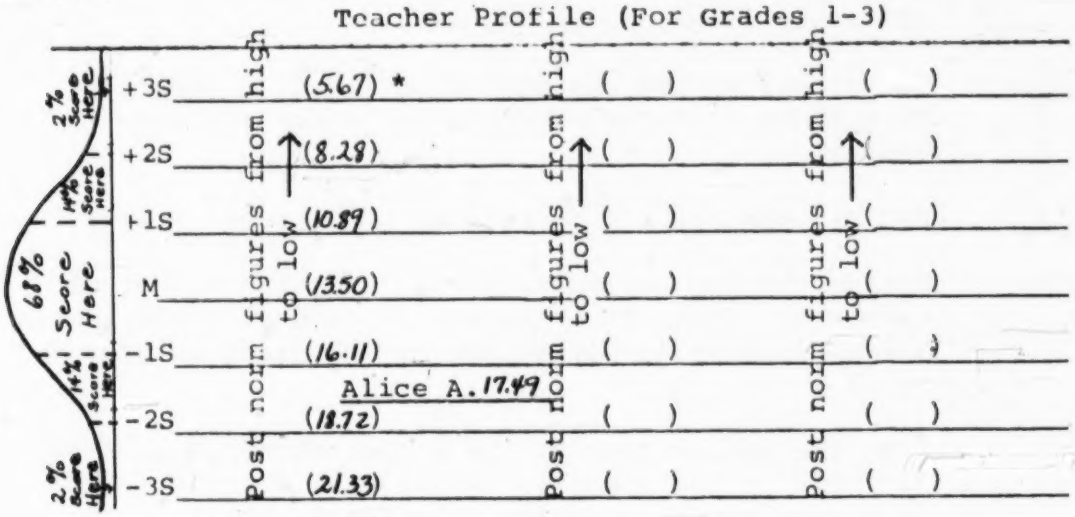
The general formula for finding the standard deviation is:

$$S.D. = \sqrt{\frac{\Sigma x^2}{N}}$$

where S.D. stands for standard deviation, Σ means "the sum of," x^2 means the squared deviations, and N is the number of teachers.

When the group mean and group standard deviation are completed for Scale 1, repeat the process for the other scales. What results will be the norms needed to post upon the Teacher Profile (see Appendix H), also as in the following illustration: Alice A.

Teacher Profile (For Grades 1-3)



	Scale: Stimulating, Interactive Style	Unreasonable Negativity	Fosterance of Self-Esteem
Group Mean	13.50	____	____
Standard Deviation	2.61	____	____

(1) Post the group mean score.
(2) Add the standard deviation to the mean score one time for 1 S.D., two times for 2 S.D., three times for 3 S.D.
(3) Subtract the standard deviation from the mean score once for -1 S.D., twice for -2 S.D., and three times for -3 S.D.
(4) Post the teacher's <u>own</u> factor scores on the chart.

The Teacher Profile is now ready for use in counseling or feedback activities.

An example of six such teacher profiles, from teachers who were included in the present sample, is presented in Appendix M.

APPENDIX L

Form for Use with Grades Above 4*

Set II (4+)

Your Name _____ Date _____

Your Teacher _____ School _____

YOUR TEACHER

Do you really notice how your teacher acts?
Please mark the following sentences about your teacher. Tell if each sentence is true or false by putting an X under T or F. Be Honest. Your teacher will NOT see these answers.

True: T (Use an X)	False: F	
( )	( )	She teaches us a lot.
( )	( )	The kids like her.
( )	( )	She listens to what we want.
( )	( )	She makes what we learn interesting.
( )	( )	She always picks on people.
( )	( )	She helps us a lot.
( )	( )	She gives us too much work.
( )	( )	We can tell how she wants things done.
( )	( )	She gets mad a lot.
( )	( )	She makes school fun.
( )	( )	She likes to teach.
( )	( )	She thinks that kids are good.
( )	( )	She thinks I work hard.
( )	( )	She thinks I have good ideas.
( )	( )	She likes us kids.
( )	( )	She likes me.
( )	( )	She likes for me to help her.
( )	( )	She thinks I can do a lot on my own.
( )	( )	She thinks I am lazy.
( )	( )	She thinks I act ugly.
( )	( )	She thinks we are a smart class.
( )	( )	She thinks I am smart.
( )	( )	She is nice when we make mistakes.

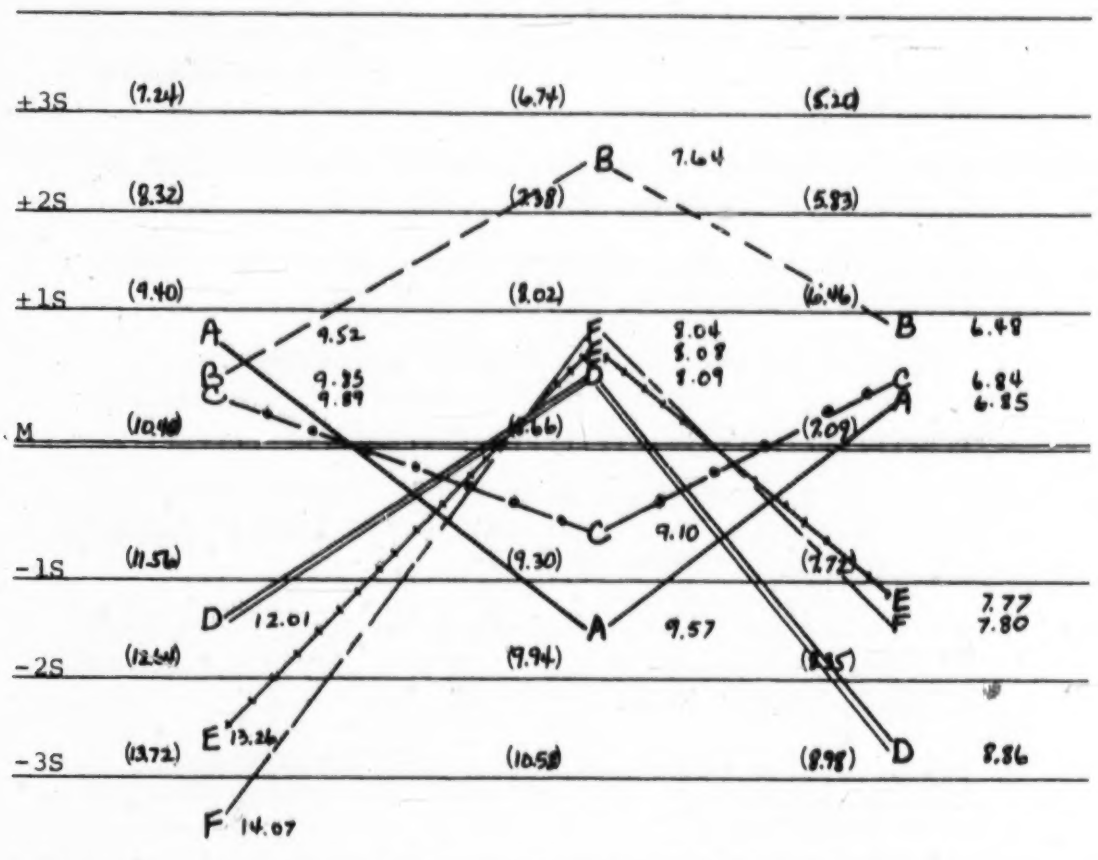
*This form substitutes for the card sorting format used below grade 4. This form should still be read aloud by the examiner when used with any group which may include a number of poor readers.

Appendix M

A COMPARISON OF SIX TEACHERS' PROFILES

(Members of This Sample)

TEACHER PROFILE



Scale 1: STIMULATING, INTERACTIVE STYLE	Scale 2: UNREASONABLE NEGATIVITY	Scale 3: FOSTERANCE OF SELF-ESTEEM

Key:

Teacher A - Sixth grade teacher
B - Second grade teacher
C - Fourth grade teacher
D - Sixth grade teacher
E - First grade teacher
F - Fourth grade teacher

REFERENCES

Amatora, Sister M. A diagnostic teacher-rating scale. Journal of Education, 1950, 30, 395-399.

Amatora, Sister M. Can elementary school children discriminate certain traits in their teachers? Child Development, 1952, 23, 75-80.

Amatora, Sister M. Teacher rating by younger pupils. Journal of Teacher Education, 1954, 5, 149-152.

Barbe, W., & Steiert, K. What children like about their teacher. Childhood Education, 1964, 40, 463-464.

Beck, W. R. Pupils' perceptions of teacher merit: A factor analysis of five postulated dimensions. Journal of Education Research, 1967, 61 (3), 127-128.

Biber, B., & Lewis, C., "An experiment study of what young children expect from their teachers." Genetic Psychology Monographs, 1949, 40, 3-97.

Bonney, M. E. Social behavior differences between second grade children in high and low sociometric status. Journal of Education Research, 1955, 48, 481.

Bower, E. Early identification of emotionally handicapped children in school. Springfield, Illinois: Charles C. Thomas, 1969.

Boyce, R. B., & Bryan, R. C. To what extent do pupils' opinions of teachers change in later years? Journal of Education Research, 1944, 37, 698-705.

Bryan, R. C. Teacher's image is stubbornly stable. Clearing House, 1966, 40, 459-461.

Burkard, M. I. Discernment of teacher characteristics by TAT sequence analysis. Journal of Education, 1962, 53 (6), 279-287.

Burkard, M. I. Effectiveness of the MTAI in a parochial school setting. Journal of Experimenal Education, 1965 33 (3), 225-229.

Cannon, K. L. Stability of sociometric scores of high school students. Journal of Education Research, 1958, 52, 43.

Christensen, C. M. Relationship between pupil achievement, pupil affect-need, teacher warmth, and teacher permissiveness. Journal of Educational Psychology, 1960, 51, 169-174.

Cohen, S. R. Exploratory study of young children's perceptions of teacher behavior. Proceedings of 75th American Convention of American Psychological Association, 1967, 2, 331-332.

Conn, L. K., Edwards, C. N., Rosenthal, R., & Crowne, D. Perception of emotional and response to teachers' expectancy by elementary school children. Psychological Reports, 1968, 22, 27-34.

Cox, F. N. Assessment of children's attitudes toward parent figures. Child Development, 1962, 33, 821-830.

Davis, O. L., Jr., & Slobodian, J. J. Teacher behavior toward boys and girls during first grade reading instruction. American Education Research Journal, 1967, 4 (3), 261-270.

Debus, R. L. Factors related to the favorability of children's perceptions of their teachers. Dissertation Abstracts, 1957, 16, 280-281.

DiVesta, F. J., & Dick, W. The test-retest reliability of children's ratings on the semantic differntial. Educational and Psychological Measurement, 1966, 26, 605-616.

Dunn, J. A., Bloom, R., & Morse, W. C. An analysis of classroom behavior from four theoretical orientations. Paper presented to Annual Meeting of the American Psychological Association, Philadelphia, 1963, 18, 358.

Ferguson, H., & Hovde, H. O. Improving teaching personality by pupil rating. School Review, 1942, 50, 439-443.

Fine, H. J., Fulkerson, S. C., & Phillips, L. Maladjustment and social attainment. Journal of Abnormal Social Psychology, 1955, 50, 33.

Gage, M. L., et. al. Teachers' understanding of their pupils and pupils' ratings of their teachers. Psychological , Monographs, 1955, 21 (Whole No. 406).

Gregerson, G. F., & Travers, R. M. W. A study of the child's concept of the teacher. The Journal of Education Research, 1968, 61 (7), 324-327.

Gronlund, N. E., & Holmlund, W. S. The value of elementary school sociometric status scores for predicting pupils' adjustment in high school. Educational Administration and Supervision, 1958, 44, 255-260.

Gronlund, N. E., & Whitney, A. P. Relation between pupils' social acceptability in the classroom, in the school, and in the neighborhood. School Review, 1956, 64, 267.

Guilford, J. P. Fundamental statistics in psychology and education. New York: McGraw-Hill Book Co., 1965.

Kearney, N. C., & Rocchio, P. D. Relation between a teacher attitude inventory and pupils' ratings of teachers, School Report, 1955, 63, 443-445.

Kratz, R. G. Characteristics of the best teacher as recognized by children. Pedagogical Seminary, 1896, 3, 413-418.

Kuhlen, R. G., & Collister, E. G. Sociometric status of sixth and ninth graders who failed to finish high school. Educational and Psychological Measurement, 1952, 12, 632.

Leeds, C. H. Teacher behavior liked and disliked by pupils. Education, 1954, 75, 29-37.

McClain, E., & Bown, O. H. Personality characteristics of student teachers whose self-descriptions differ markedly from descriptions by their pupils. American Digest, 1961, 16, 403.

McKeachie, W. J. Motivation, teaching methods, and college learning. In M. R. Jones (Ed.), 1961 Nebraska Symposium on Motivation. Lincoln: University of Nebraska Press, 1961, 111-146.

McKeachie, W. J., Isaacson, R. L., & Milholland, J. E. Correlates of teacher personality variable and student ratings. Journal of Educational Psychology, 1963, 54 (2), 110-117.

Neale, D. C., & Proshek, J. M. School-related attitudes of culturally disadvantaged elementary school children. Journal of Educational Psychology, 1967, 58 (4), 238-244.

Ojemann, R. H., & Snider, B. C. The development of the child's conception of the teacher. The Journal of Experimental Education, 1963, 32 (1), 73-77.

Patton, R. A., & DeSena, P. A. Identification through student opinion of motivating and non-motivating qualities of teachers. Journal of Teacher Education, 1966, 17, 41-45.

Paraskevopoulos, I. How students rate their teachers. Journal of Education Research, 1968, 62 (1), 25-29.

Phillips, B. N., & DeVault, M. V. Relation of positive and negative sociometric valuations to social and personal adjustment of school children. Journal of Applied Psychology, 1955, 39, 409.

Pittman, J. A. A study of the suitability of an attitude-gauging instrument for assaying the attitude-toward-schooling of a group of sixth grade Negro pupils. Journal of Negro Education, 1952, 2, 136-147.

Reed, H. B., Jr. Teacher variances of warmth, demand, and utilization of instrinsic motivation related to pupils science interests: A study illustrating several potentials of variance-covariance. Journal of Experimental Education, 1961, 29, 205-209.

Ryans, D. G. Characteristics of teachers. American Council on Education, 1960, 397-398.

Spillard, Sister M. A. The tale of twenty tattlers. Grade Teacher, March, 1964, 81, 116-118.

Symonds, P. M. Characteristics of the effective teacher based on pupil evaluations. Journal of Experimental Education, June, 1955, 23, 289-310.

Taylor, P. H. Children's evaluations of the characteristics of the good teacher. British Journal of Educational Psychology, November, 1962, 32, 258-266.

Tenenbaum, S. Attitudes of elementary school children to school, teachers, and classmates. Journal of Applied Psychology, 1944, 28, 134-141.

Tuckman, B. W., & Oliver, W. F. Effectiveness of feedback to teachers as a function of source. Journal of Educational Psychology, 1968, 59, (4), 297-301.

Ullman, C. Teachers, peers and tests as predictors of adjustment. Journal of Educational Psychology, 1957, 48, 257-267.

Veldman, D. J. Pupil evaluation of student teachers and their supervisors. Journal of Teacher Education, Summer, 1970, 21 (2).

Veldman, D. J., & Peck, R. F. Student teacher characteristics from pupils' viewpoint. Journal of Educational Psychology, 1963, 54, 346-355.

Weintraub, S. Pupil conceptions of the teacher. The Reading Teacher, 1967, 20, 441-446.

White, W. F., & Anderson, H. E. A study of scaled dimensions of teacher behavior as perceived by students. Journal of Psychology, 1967, 65 (2), 223-232.

White, W. F., & Dekle, O. T. Effects of teacher motivational cues on achievement level in elementary grades. Psychological Reports, 1966, 18, 351-356.

Wilson, L. A. Do children know who are the best teachers? Educational Forum, 1948, 13, 63-67.

Wright, B., & Sherman, B. Love and mastery in the child's image of the teacher. The School Review, Summer, 1965, 73, 89-101.

SET2 COMPUTER PROGRAM MANUAL

by

Bob Pennington

The SET2 program was written to provide detailed feedback concerning students' evaluation of their teacher. Data used by the SET2 program is provided by the Student Evaluation of Teacher Instrument .II (SET .II), developed by Haak, Kleiber, and Peck of the Research and Development Center for Teacher Education. A copy of the SET .II manual may be obtained by contacting one of the authors.

The SET .II contains 22items, twenty of which comprise three scale scores and two subscale scores. The items which cor spond to each scale are:

Rapport Subscore	5, 4, 16, 13, 18
Interactional Competence Subscore	7, 9, 3, 17
Stimulating, Interactive Style Scale Score	R Subscore + IC Subscore
Unreasonable Negativity Scale Score	8, 22, 20, 10, 6
Fosterance of Self-Esteem Scale Score	15, 12, 14, 11, 19, 21

INPUT

Header Card. A Header card is required for each teacher. The card includes the number of students participating in the evaluation and a sixty column space usually used to convey or record information or identification of the teacher. Different classes may be run for each teacher sequentially but a header card is necessary for each class. The format for the header card is (I5,10A6).

Data Card. There is one data card for each student. The first forty columns are not read by the SET2 program and can thus be used to identify students in any number of ways. Columns forty-one through sixty contain twenty one-digit SET .II scores. These scores are either a "1" or a "2" indicating "true" or "false" respectively. The data card format is (40X,20F1).

Deck Arrangement. The header card followed by the appropriate number of data cards constitutes a set of data. Any number of sets can be processed under the restraints of time and paper allocation. The last set of data processed must be followed by a blank card. When completed the deck should resemble the following:

Header Card for Teacher
Data Cards of Teacher's students
.
.
.
Header Card for Teacher
Data Card for Teacher's students
.
.
.
Blank Card.

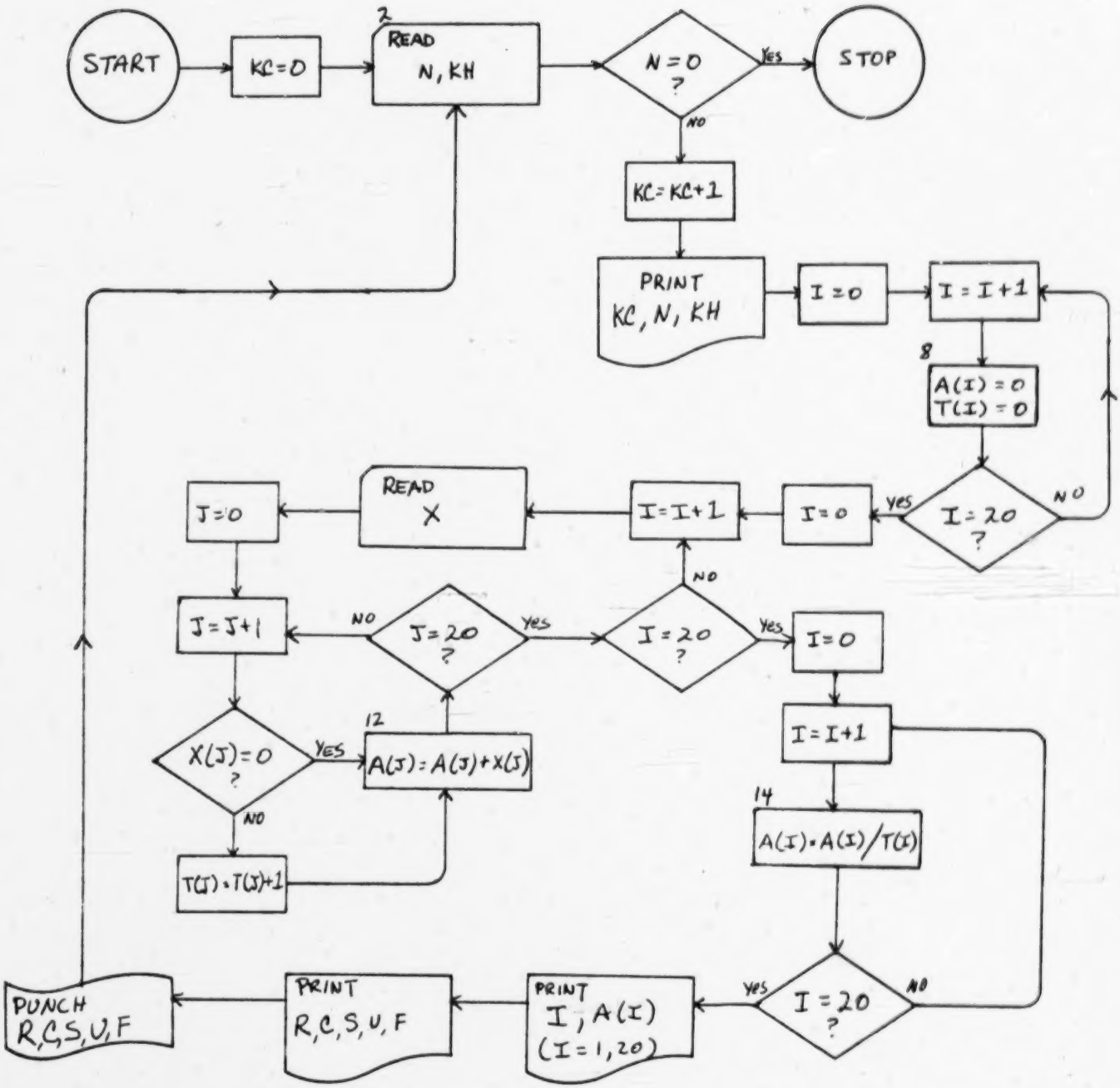
OUTPUT

The header card (KH) is outputed along with the class number (KC) and the number of students in the class (N). After the item means have been tabulated they are outputed vertically with the mean for each item following the item number. The three scale scores and the two subscores are then computed and outputed following the item means

Appended to this manual are copies of the program listing, program flowchart, and program output. The author wishes to express his appreciation to Dr. Doug Kleiber who developed the original SET2 program.

```
      PROGRAM SET2 (INPUT,OUTPUT,PUNCH)
      DIMENSION A(20),X(20),T(20),KH(10)
      KC=0
    2 READ 4, N, KH
    4 FORMAT (I5, 10A6)
      IF (N.EQ.0)   STOP
      KC = KC + 1
      PRINT 6, KC, N, KH
    6 FORMAT(*1ITEM *,*MEANS *,*FOR CL*,*ASS*,I3,* N=*,I3,2X,10A6/)
      DO 8 I = 1,20
    8 A(I) = T(I) = 0
      DO 12 I=1,N
      READ 10,X
   10 FORMAT (40X,20F1)
      DO 12 J = 1,20
      IF (X(J).GT.0) T(J) = T(J) + 1
   12 A(J) = A(J) + X(J)
      DO 14 I=1,20
   14 A(I) = A(I) / T(I)
      PRINT 16,(I,A(I), I=1,20)
   16 FORMAT (I5,F10.4)
      R=C=S=U=F=0
      R = A(1)+A(2)+A(3)+A(4)+A(5)
      C = A(6)+A(7)+A(8)+A(9)
      S = R + C
      U = A(10)+A(11)+A(12)+A(13)+A(14)
      F = A(15)+A(16)+A(17)+A(18)+A(19)+A(20)
      PRINT 22,R,C,S,U,F
   22 FORMAT (//,1X,*SUBSCO*,*RE RAP*,*PORT =*,* *,F8.2,/,1X,*SUBSCO*,
     1*RE INT*,*ERACTI*,*ONAL C*,*OMPETE*,*NCE = *,F8.2,/,1X,*STIMUL*,
     2*ATING,*,* INTER*,*ACTIVE*,* STYLE*,* SCALE*,* SCORE*,* = *,F8.2,
     3/,1X,*UNREAS*,*ONABLE*,* NEGAT*,*IVITY *,*SCALE *,*SCORE *,*= *,
     4F8.2,/,1X,*FOSTER*,*ANCE O*,*F SELF*,*-ESTEE*,*M SCAL*,*E SCOR*,
     5*E = *,F8.5)
      PUNCH 24,KC,R,C,S,U,F
   24 FORMAT (1X,*C*,I2,7X,5F8.2)
      GO TO 2 $ END
```

Program Flowchart SET2



Program SET2 _Output

```
ITEM MEANS FOR CLASS  1 N=  5        SAMPLE DATA FOR SET 2   WRITTEN 3/73   PENNINGTON

    1     1.4000
    2     1.4000
    3     1.4000
    4     1.8000
    5     1.2000
    6     1.8000
    7     1.8000
    8     1.4000
    9     1.8000
   10     1.6000
   11     1.8000
   12     1.6000
   13     2.0000
   14     1.6000
   15     1.8000
   16     1.2000
   17     1.4000
   18     1.4000
   19     1.2000
   20     1.0000


SUBSCORE RAPPORT =      7.20
SUBSCORE INTERACTIONAL COMPETENCE =      6.80
STIMULATING, INTERACTIVE STYLE SCALE SCORE =     14.00
UNREASONABLE NEGATIVITY SCALE SCORE =      8.60
FOSTERANCE OF SELF-ESTEEM SCALE SCORE =  8.00000
```